COMPETITION PROCESSES IN VISUAL WORD RECOGNITION

JULIE KATE VOICE

DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF EDINBURGH

1995

For Sandy

Declaration

I declare that this thesis has been composed by myself and that the research reported therein has been conducted by myself.

London, 1 July 1995

Julie Kate Voice

Acknowledgments

Thanks to my supervisors, Ellen Bard and Richard Shillcock, for the ideas, advice, guidance and encouragement they have given me. I am especially grateful to Richard for helping me keep things in perspective and stay reasonably sane.

Thank you to Lolly and William for giving me the time and space I needed to finish this thesis, while working in their lab.

I wish to thank Mark Seidenberg for very generously giving me a copy of the unpublished item means from his Mega Study.

I am very grateful to Steve Finch of the H.C.R.C. in Edinburgh for writing a neighbourhood finding program, and for providing letter bigram frequencies. I am also grateful to Steve Barton of Glasgow University for running the eye-tracking experiment reported in Chapter 5.

A big thanks to everyone who's listened to my thesis moans and worries over the last few years, everyone who's been so supportive in times of stress and despair. Thank you especially, to Mum, Dad, and Bean at home in Somerset, to Diane, Louise, Simon, David P, Dave, Dan and Karen in Edinburgh, and to Sam and Mark in London.

Abstract

In this thesis we show that lexical decisions (LDT) to isolated words are slowed when a target word has either, many orthographic competitors (defined as words sharing outer letter frames), or a single strong competitor (defined as words sharing all letters, e.g. from/form). This result is not found for naming latencies. It is shown that naming is more sensitive to variables which concern the mapping of orthography to phonology, than to purely orthographic measures. In contrast to these findings Andrews (1989, 1992) found that the existence of orthographic competitors in the form of neighbours (Coltheart, Davelaar, Jonasson and Besner (1977), facilitated both LDT and naming responses. This suggests that neighbours conspire rather than compete for recognition. We repeat Andrews' experiments and replicate the conspiracy effect of neighbourhood size for low frequency words. Andrews' results replicate however, only for LDT and not for naming. In a further investigation of neighbourhood effects we measure eye fixations to target words in neutral sentence contexts. Here we find an effect of neighbourhood size for both high and low frequency words. The effect is, however, one of conspiracy for high frequency words, and one of competition for low frequency words. We describe an activation time course model of word recognition which we argue can account for both competition and conspiracy effects, and for the differences in effects for high and low frequency words.

Neighbourhood influences are less robust for naming than for LDT or sentence reading. This may be because isolated word naming relies less on lexical processing than the other two tasks. Whether naming responses are produced lexically, or by strict grapheme-phoneme correspondence rules, will depend on the proportion of spelling-sound regular and irregular words in the stimulus list. We demonstrate the influence of stimulus list structure on naming by comparing data from the Seidenberg and Waters (1989) Mega Study with data from

smaller studies. For tasks which demand more lexical processing, such as LDT, we demonstrate that the influence of stimulus structure will depend on the disparity between lexical knowledge and local information obtained from the stimulus list. Stimulus list structure is most influential when it gives little, or no, reflection of real lexical experience.

We discuss the implications of all our findings in terms of modeling the process of visual word recognition, and in terms of current and future experimental research.

Table of Contents

1. Introduction **1**

1.1 The word superiority effect 2

1.2 The word frequency effect 4

1.3 The regularity effect 5

1.4 The influence of candidate sets 6

1.5 The influence of candidate sets in models of lexical access 7

1.6 Structure of the thesis 9

2. Literature Review **14**

2.1 Competition processes in lexical access 14

2.1.1 Outline shapes and envelopes 15

2.1.2 Single letters 16

2.1.3 Letter clusters 18

2.1.4 Whole words 20

2.1.5 Summary: competition effects 22

2.2 Lexical and sublexical effects 24

2.2.1 LDT and Naming 25

2.2.2 Lexical and sub-lexical naming 26

2.2.3 Pronunciation influences in LDT 27

2.2.4 Models of word naming 28

2.2.5 Summary: lexical and sublexical effects 30

2.3 Strategies in visual word recognition 31

2.3.1 Strategies 31

2.3.2 Stimulus lists 33

2.3.3 Summary: strategies in visual word recognition 36

3. General Methodology **37**

3.1 Stimulus specifications 37

3.1.1 Word length 38

3.1.2 Word frequency 38

3.1.3 Orthography 39

3.1.4 Concreteness 41

3.1.5 Word onset 43

3.1.6 Nonwords 44

3.2 Tasks 45

3.2.1 Lexical decision 45

3.2.2 Naming 46

3.2.3 Sentence reading 47

4. Competition effects **48**

4.1 The effect of a single strong competitor 48

4.1.1 Experiment 1a: LDT 50

4.1.2 Results 53

4.1.3 Experiment 1b: Immediate naming 54

4.1.4 Results 55

4.1.5 Summary of Experiment 1 55

4.2 The effect of competitor set size 55

4.2.1 Experiment 2a: LDT 56

4.2.2 Results 59

4.2.3 Experiment 2b: Immediate naming 60

4.2.4 Results 61

4.2.5 Summary of Experiment 2 61

4.3 Competition at word onset 62

4.3.1 Experiment 3 63

4.3.2 Results 67

4.3.3 Summary of Experiment 3 68

4.4 Discussion 69

4.4.1 Competition between candidates 69

4.4.2 Conspiracy between candidates 70

4.4.3 Position Specificity 72

5. Orthographic neighbourhood size **73**

5.1 Orthographic neighbourhood size effects in LDT and naming . . 73

5.1.1 Experiment 4a: LDT 74

5.1.2 Results 79

5.1.3 Experiment 4b: Immediate and delayed naming 82

5.1.4 Results 83

5.1.5 Summary of Experiment 4 85

5.2 Connectionist models of the neighbourhood effect 86

5.2.1 Simulation 1 87

5.3 Orthographic neighbourhood size effects in reading 89

5.3.1 Results 91

5.3.2 Summary for Experiment 5 96

5.4 Discussion 97

5.4.1 Time course of activation 98

5.4.2 Task differences 103

6. Stimulus list effects **107**

6.1 Effect of stimulus list context in naming 107

6.1.1 The Seidenberg and Waters (1989) Mega Study 108

6.1.2 Results 109

6.1.3 Summary 112

6.2 Effect of Stimulus list context in LDT 113

6.2.1 Results 120

6.2.2 Summary of Experiment 7 121

6.3 Blocked presentation of orthographic neighbourhood size 122

6.3.1 Experiment 8a: LDT 123

6.3.2 Results 124

6.3.3 Experiment 8b: Immediate naming 127

6.3.4 Results 128

6.3.5 Summary of Experiment 8 128

6.4 Discussion 130

6.4.1 Stimulus list influence in word recognition 130

6.4.2 Influential stimulus list information 132

6.4.3 Stimulus list and neighbourhood effects 133

7. Discussion and Conclusions 135

7.1 Implications for models of isolated word recognition 135

7.1.1 Visual encoding and position specificity 135

7.1.2 Lexical access and competition 136

7.1.3 Reverberation of activation and conspiracy 137

7.1.4 Processing routines: one or two? 138

7.1.5 Top-down contextual influences 139

7.2 Implications for other research 140

7.2.1 Priming paradigms 140

7.3 Defining candidate sets 145

7.4 Summary of Thesis findings 147

7.4.1 Do candidates compete or conspire? 147

7.4.2 How does word frequency influence competition or conspiracy? 148

7.4.3 Why are neighbourhood effects weaker for naming than for LDT? 148

7.4.4 How can bigrams be independent of neighbourhoods? . . 149

7.4.5 What determines the influence of stimulus list? 149

A. Stimuli for Experiment 1 **170**

B. Stimuli for Experiment 2 **171**

C. Stimuli for Experiment 3 **172**

D. Stimuli for Experiments 4 and 8 **174**

E. Stimuli for Experiment 5 **175**

F. Mega study comparisons **177**

G. Stimuli for Experiment 7 **187**

H. Relatedness within a random sample of 1625 monosyllabic words 188

Chapter 1

Introduction

Although the typical adult vocabulary contains about 50,000 words (Monsell, Doyle and Haggard 1989), a skilled reader is able to respond to an isolated word in under half a second. It is extremely unlikely that the reader has to consider all 50,000 possible lexical entries each time a single word is recognised. It is more probable that just a subset of the lexicon, a candidate set of potentially correct words, is considered. A single correct item must then be selected from these candidates. The metaphor used to describe the selection process is that of competition, the members of a candidate set compete until one item wins and is thus recognised. The issues of how a candidate set is defined, and how members of a candidate set influence one another, are very controversial. These issues are what we are concerned with in the present thesis.

Before we begin to consider candidate sets and competition processes, let us consider a number of relatively less controversial effects which we observe in skilled reading, and that any model of the basic reading process must explain. In the first three sections of this Chapter we describe these effects and how they are accommodated in models of word recognition. The aim of these three sections is to set out what we know about the recognition process, and to introduce the language and metaphors we use to discuss it. In the fourth section, we then turn to the more controversial issue of candidate set influence. We consider how well models accommodate any effect of non-target items on target recognition. In the final section of this Chapter we describe the structure and content of the remaining six thesis chapters.

1.1 The word superiority effect

Skilled readers can distinguish real words from nonword letter strings very quickly, even when the nonword strings are very wordlike. The speed with which readers recognise that a word is a word (about 600 to 700ms) is faster than the speed with which they recognise that a nonword string is not a word (about 800 to 900ms). Single letter probe recognition studies have also demonstrated word superiority: letters within words are more quickly identified than letters in nonword strings (McClelland and Johnston 1977; Masson 1975; Massaro, Venexky and Taylor 1979).

One of the simplest ways to represent the word/nonword distinction in a model of lexical access is as a failure of nonword strings to be represented at a lexical level. Nonwords are letter strings we cannot find when we look them up in our lexicon. This is the sort of description used in the serial search model (Rubenstien, Garfield and Millikan 1970; Rubenstien, Lewis and Rubenstien 1971; Forster 1976). The lexicon, or sections of it, are serially scanned, item by item, until a match between the sensory information and a lexical representation is found. A search for a nonword string will simply fail to recover a lexical entry. Nonwords take longer to be rejected than real words do to be accepted because a search for a real word is terminated once that word is found, whereas a nonword search must always be exhaustive. The serial search model does not accommodate the effects of different sorts of nonwords, a string is either in the lexicon or it is not. There is evidence to suggest that the wordlikeness of nonwords' structure has some influence on processing; this can be better accounted for in models which consider the metaphor of activation.

The earliest activation based model of word recognition is the logogen model (Morton 1969, 1970) which has entries in the mental lexicon represented as logogens. These are evidence collecting devices which become increasingly active as they match incoming sensory information. The more features a logogen shares with a sensory stimulus, the more active it will become. Every logogen has a predetermined evidence threshold and when the sensory information provides enough activation to exceed the threshold, the logogen will fire. Once a logogen has fired, the word it represents becomes available as a response and can be recognised. A nonword string may provide some matching sensory information and thus activation to some logogens, but it will never cause one to fire. For this reason, nonword rejection is slow: responding is delayed while any logogens are active and there is the chance that one may exceed threshold and fire. The more wordlike a nonword is, the more logogens it will activate and thus the longer it will take to reject.

The Interactive Activation Model (McClelland and Rumelhart 1981; Rumelhart and McClelland 1982) offers a similar account. Instead of logogens the model consists of processing units which correspond to various levels of language structure, these are visual or auditory feature units, letter or phoneme units and word units. On presentation of a written word visual feature units are activated which in turn activate appropriate letter units, which go on to activate word level units. Activation feeds forward though levels in this manner, but as it builds up in higher levels it feeds back down to influence the continued processing of lower level units, hence the system is interactive. Presentation of a nonword string will activate featural and letter level units which will in turn pass activation to word level units, but no word level unit will receive enough activation to result in recognition.

Connectionist models of word recognition have developed from the principles of interactive activation. Unlike earlier activation or search models however, connectionist architectures, such as that of Seidenberg and McClelland (1989) have no explicit lexicon or word representations. The Seidenberg and McClelland model consists of an associative network of nodes and connections. Activation flows through the connections to activate nodes. Unlike the Interactive Activation Model the nodes do not correspond to specific letters, phonemes, or words. Instead, specific lexical items are represented as patterns of activation distributed across the entire network. The model has a set of orthographic units, a set of phonological units and a set of hidden units which mediate between the orthographic and phonemic units. During a training phase the model is exposed to a large number of words. Different words produce different patterns of activation across the network and affect the connection weights. By experiencing many words the model learns to encode generalisations about lexical structure. Although nonwords will give rise to patterns of activation never before experienced by the model, the more wordlike the nonword is the more likely the model is to either mistake it for a word it knows, or to take a long time to reject it.

1.2 The word frequency effect

Words with a high printed frequency of occurrence are recognised more quickly than words with lower frequencies (Andrews 1989, 1992; McRae, Jared and Seidenberg 1990; Segui and Grainger 1990).

How a model of lexical access explains the word frequency effect depends on whether the adopted metaphor is one of search or activation.

In the serial search model all searching is frequency biased. The most frequent lexical items will be those scanned first and recognised fastest.

Within activation frameworks the word frequency effect is represented by evidence thresholds (Morton 1969, 1970), or by resting activation levels (McClelland and Rumelhart 1981; Rumelhart and McClelland 1982). High frequency words have lower evidence thresholds, or greater amounts of resting activation, than low frequency words. The result is that less evidence has to accumulate before a high frequency word can be recognised, and therefore high frequency words are recognised more quickly than low frequency words. Seidenberg and McClelland's (1989) model simulates frequency effects because more frequent words are presented to the model more often during its training phase. Patterns of activation which are experienced most often have the greatest impact on connection weights, and thus, those patterns become easier for the model to recognise.

1.3 The regularity effect

Skilled readers are able to pronounce many words, but are also able to produce pronunciations for nonword letter strings which they have never seen before. Word pronunciation is possible whether the mapping of letters to sounds is regular (e.g. *bead, mead, read*), or irregular (e.g. *dead, head*). Typically however, irregular words have longer response latencies than regular words (Gough and Coskey 1977; Stanovich and Bauer 1978; Parkin and Underwood 1983).

Explanations of the regularity effect and pronunciation of nonwords vary depending on whether a model has one or two processing "routes".

Dual route models (Coltheart 1978, 1985) have two different procedures by which print is converted to a phonological code. These are a dictionary type lookup of whole words and a letter-to-sound rule-based procedure. The lexical look up route accesses stored pronunciations. The fact that people can pronounce letter strings they have never seen before may be evidence that additional to lexical look-up, there must also be a nonlexical system of rules specifying the relationships between letters and sounds in English. Both lexical and nonlexical routes provide correct pronunciations for regular words. The nonlexical route will provide correct pronunciations only for words obeying the spelling-sound rules of English; it will give incorrect pronunciations for irregular words. Regular words will produce the same output from both routes whereas irregular words may result in different outputs from the two routes. Resolution of this mismatch will delay the recognition of irregular words.

In contrast, the connectionist architecture of Seidenberg and McClelland (1989) explains pronunciation and regularity effects with only a single processing mechanism. During the model's training phase it is exposed to a large number of orthographic/phonological pairs (e.g.*keen/ki:n*) and the desired pattern of activation within the mediating hidden units evolves via a back propagation learning algorithm. By experiencing many words the weighted connections of the network encode the correlations among word variables (e.g. spelling-sound rules). A stimulus word producing a pattern of activation close to one on which the model has been trained, will yield the correct orthographic or phonemic output. The preferred pronunciation for any post-training presented letter string will depend on the highest output score derived from the pattern of activation in the hidden units. Words that share orthography and phonology (e.g.*bead, mead*) have similar effects on connection weights and so exposure to one improves the other. The presence of words sharing orthography but having different phonology (e.g. *bead, dead*) means that training on one will have a negative effect on the other. The result is poorer performance on irregular words.

1.4 The influence of candidate sets

If nothing at all was known about a stimulus, except that it was a word, then every lexical entry would be a possible competitor. It is thought, however, that early sensory information narrows the candidate set so that only a small portion of the lexicon is considered. The nature of word candidate sets and how they affect target words is controversial. Most researchers agree however that the identification of a single word is in some way influenced by other members of its candidate set. Potential sources for the influence of non-target candidates include candidate set size, i.e. the number of competitors a word has (Andrews 1989, 1992; Forster and Davis 1984; Colombo 1986), and the frequency of candidate set members relative to target frequency (Grainger, O'Regan, Jacobs and Segui 1989; Segui and Grainger 1990). The effect of candidate set members on target word recognition has been reported as being facilitatory (Andrews 1989, 1992), and as being inhibitory (Coltheart 1977). In the next section we consider how models of lexical access can accommodate these sources and types of candidate set influence.

1.5 The influence of candidate sets in models of lexical access

In a serial search framework it would be nonsensical to suppose that the entire lexicon was represented as a word list and that each time a string of letters was perceived a search began at the top of the list and ran down until a match was made. Instead each search is considered to be confined to what Forster (1976) has called a "bin". The bins represent candidate sets and are defined by some predetermined sensory characteristics. For example all words beginning with "*ph*" may constitute a bin. There is no specification of what can or cannot define a bin, but bins are searched in a frequency order. A word may therefore be influenced by the frequencies of non-target candidates. The more frequent a candidate is relative to the target, the more likely it is that the candidate will have to be checked and rejected before the target is scanned and correctly identified. Word recognition will be slower if a word's candidate set contains many high frequency words because there will be more interruptions for checking during the search process.

Because logogens are independent and unordered, the logogen model has problems accounting for any finding which involves the recognition of one word being influenced by the presence of other words. Once a stimulus is present and logogens are active, there is a race to see which will reach threshold first. The thresholds of logogens are fixed and so cannot be affected by the presence of other active logogens. Whether a stimulus activates 2 or 200 logogens, the target will have the same activation threshold to exceed.

In contrast, Interactive Activation Models can explain the influence of candidate set members. The more candidates a target activates, the more activation there will be feeding back to the constituent letter or phoneme units of the target word. Words from large candidate sets will therefore be more quickly recognised. Interactive activation models can also predict exactly the opposite effect of candidates, that they are inhibitory to target word recognition. Active non target word representations will laterally inhibit the target and thus the more active candidates in competition, the more inhibition a target will receive and the more slowly it will be recognised. The inhibitory output of a given node increases as its activation level rises and is therefore a function of its resting potential which derives from its word frequency. High frequency candidates will therefore provide stronger inhibition effects on the target than lower frequency candidates. Like the serial search model, an interactive activation account predicts slower recognition for a target with high frequency candidates than for one with less frequent set members.

The Seidenberg and McClelland (1989) model supports similar predictions to those of the Interactive activation model. The larger a word's candidate set, the more similar words there are, and thus, the more likely it is that the activation pattern of the target will provide at least partial activation of many other words. This may mean that the target is harder to discriminate, available resources are more likely to be widely distributed across the system than in the case of a small candidate set. If candidates sets contain words that share featural overlap, then large sets will consist of word with common feature sets. These features will have been experienced more often by the model and this will be reflected in the connection strengths. A word with many candidates may therefore be easier to recognise.

1.6 Structure of the thesis

The next two chapters of this thesis offer further background information for the later experimental work. Chapter 2 presents a review of the literature considered to be most relevant to issues dealt with in the thesis. The review is split into three major sections which deal with, competition processes in lexical access, lexical and sublexical processing, and the use of strategies in lexical processing. After each of these sections we summarise the current evidence and set out the questions to be addressed by experimental investigations later in the thesis. Chapter 3 introduces the general methodological issues of concern to the thesis work. We consider details of the response tasks, variables, and controls, used in the experimental investigations.

From the literature reviewed in Chapter 2, it is apparent that while word parts may conspire, it is unclear as to whether complete lexical representations also conspire or instead compete. In Chapter 4 we present two new experiments which manipulate lexical level candidates. Experiment 1 shows that competition can be provided by a candidate which is visually similar to, and higher frequency than, the target word. Experiment 2 shows that competition is provided when a target comes from a large candidate set of possible words. We suggest that these two sources of competition are related since large candidate sets are found to have stronger competitors in terms of frequency, and thus activation advantage over a target.

In Chapter 4 we also consider the problem of position specificity. It is clear that the word recognition system has some sensitivity to letter position within words since items such as *live* and *vile* can be distinguished. The position specificity devices used in models of word recognition are often, however, very restrictive. The findings of Experiment 1 demonstrate the inadequacy of one such device, that of Wickelgrams, in accounting for competition between non-position specified letter sharing words.

The competition effects of candidate strength (Experiment 1) and candidate number (Experiment 2) are found only for LDT and not for word naming. We suggest this is because in both experiments the candidate sets were defined by measures of orthography. For the naming task it may be the spelling to sound mapping which is more important than pure orthography. In Experiment 3 we measure candidate sets by determining the consistency of digram to diphone correspondence at word onsets. The more words which share an onset digram to diphone mapping, the more consistent the mapping is, and the more quickly a word is named. Since onsets are, however, word parts, their conspiracy is not surprising.

Unlike word onset digram to diphone correspondence consistency, orthographic neighbours (Coltheart, Davelaar, Jonasson and Besner 1977) are whole word candidates and do not appeal to the spelling to sound mapping. Andrews (1989, 1992) finds, however, that orthographic neighbours conspire, and furthermore, that they do so in naming as well as LDT. In Experiment 4 we repeat the experiments of Andrews with new controls and are able to replicate her conspiracy effects only for LDT. We suggest this may be due to the exaggeration of neighbourhood effects in the decision process of LDT, or an overshadowing of the effect by the pronunciation process in naming.

In Experiment 5 we use eye tracking measurements to look at neighbourhood effects in a sentence reading task which involved no decision or pronunciation. We find a significant influence of neighbourhood size on eye fixation times in reading and therefore suggest that neighbourhood is important to the general lexical access process. The reason we give for neighbourhood effects being unreliable in word naming, is that naming can be performed lexically or sublexically by GPC rules, whereas LDT and reading will normally be performed lexically. We believe that only when processing is lexical will a neighbourhood effect become apparent. A single route connectionist model such as that of Seidenberg and McClelland (1989), has no explicit lexicon, but is shown in Simulation 1, to simulate the conspiracy effects of orthographic neighbourhood size. With reference to recent connectionist work we argue that multi layer architectures like the Seidenberg and McClelland model result in a form of lexical processing, that is that their hidden units come to represent a distributed lexicon. Multi layer connectionist networks are therefore very successful in modeling lexical level effects and do not pose a threat to our assumptions concerning the weaker effects in naming.

The effect of neighbourhood size for LDT in Experiment 4 is one of conspiracy for low frequency words. In contrast, Experiment 5 shows a conspiracy effect of neighbours for high frequency words but a competition effect for low frequency words. LDT responses produce much longer latencies than reading, we therefore suggest a time course mechanism to accommodate the changing neighbourhood effects. If we assume that high frequency words become active more quickly than low frequency words, then they may show the influence of orthographic neighbourhood size earlier than low frequency words. This is what we find when we compare Experiments 4 and 5. Neighbours conspire for high frequency words in reading at short latencies, and conspire for low frequency words in LDT when response latencies are much longer.

The difference in neighbourhood effects for low frequency words between the two tasks suggests that neighbours first compete and then later conspire. We suggest that the conspiracy effect is due to reverberating activation between word and lower levels of representation. The existence of a target word is reinforced by activation from its constituent parts. Reverberation is a process which takes time to operate within a system since activation must build up before it can begin to reverberate. A conspiracy effect based on reverberating activation will, therefore, always occur after competition effects.

From the reviewed literature it is suggested that the use of lexical or sub-lexical processing in word naming depends on the proportions of regular and irregular words in a stimulus set. Only when a stimulus list is made of predominantly regular words will simple grapheme to phoneme correspondence be a reliable process. In Experiment 6 we compare naming times from the Seidenberg and Waters (1989) Mega Study with naming times from smaller studies. These cross study comparisons compare responses to the same items on the same task, but show quite poor replication of results. The structure of the stimulus lists is what varies between the studies considered, and it is this source of variation which we believe is responsible for changes in responses.

For a task which uses predominantly lexical processing, such as LDT, there may be many lexical dimensions on which the stimulus list may be influential. Many word recognition effects (e.g. word frequency, orthographic redundancy) reflect lexical experience. We suggest therefore, that the relationship between the lexical knowledge a subject has, from years of reading experience, and the local information received from a particular stimulus list, may be the key to determining stimulus list influence. We believe that top-down contextual information reverberates back to lower levels and that when there is a discrepancy between this information and what the system already knows, then the stimulus list is able to provide a very potent source of influence. In Experiment 7 we demonstrate that subjects are more greatly influenced by the structure of a stimulus list when that structure is not representative of normal lexical experience. In Experiment 8 we use this finding to encourage an influence of neighbourhood size in naming. By demonstrating that neighbourhood size can influence naming, as well as LDT and reading, we gain further support for the claim that orthographic neighbourhood size is important to a general lexical access process.

In Chapter 6 we also show that dimensions of the stimulus list which can be used to operationalise the WORD/NONWORD discrimination in LDT are very influential, even when they are not a variable we would normally consider to be relevant to lexical processing.

Chapter 7 is a general discussion of the thesis findings. We consider the implications of the findings for models of word recognition, for current and future research. We also consider the nature of lexical variables and how candidate sets are defined. Finally we present a summary of the thesis in terms of the questions it has answered.

Chapter 2

Literature Review

The purpose of this Chapter is to review literature of concern to the theoretical considerations dealt with in the present thesis. The review will be divided into three major sections. After each of these we will summarize the current position, and set out what is to be investigated later in this thesis.

2.1 Competition processes in lexical access

In visual word recognition, for a target item to be correctly recognised, it must be distinguished from all other possible alternatives. The recognition process is assumed not to consider every known word each time a target is perceived, but to restrict consideration to a candidate set of potential items. Competition is thought to occur among these candidate items, and give rise to a winner, the target representation. Since words are easier to remember or read than random strings, it appears that the recognition system intelligently capitalises on some structure that exists within word arrays. There are many possible levels of representation we can employ to describe the structure of words, e.g. letters, bigrams, trigrams, syllables, rhymes or word shapes. Candidate sets may therefore consist not of whole word items, but of word parts, such as bigram sequences or syllables. The literature of competition effects can be organised by the level of representation at which competition has been considered to occur. In the following four sections we review the evidence that exists for and against the existence of competition processes at different levels of representation.

2.1.1 Outline shapes and envelopes

Navon (1977) has used the term "envelope" to refer to the outline shape of a word. The general claim is that across a variety of tasks, words with distinct envelopes are more easily recognised than equally frequent but less distinct controls (Havens and Foote 1963; Bouma 1971; Marchbank and Levin 1965; Rayner 1976; Dunn-Ranklin 1978). The more envelope sharing candidates there are, the more competition for recognition there will be. Evidence for the existence of independent word envelope effects is, however, intrinsically problematic because envelope similarity is highly correlated with individual letter similarity. Posnansky and Rayner (1977) demonstrated that picture naming was facilitated both by letter sharing and envelope sharing nonword primes. If *list* is more readily confused with *lint* than with *lilt*, we cannot tell whether this is an envelope effect, or an effect of the letter *s* being more like *n* than like *l*.

If word envelopes were fundamentally important to the recognition of a word, then lowercase words should be easier to recognise than those presented in uppercase, since lowercase letters provide more distinct envelopes, and thus will have fewer competitors. Yet McClelland (1976) found no difference in report accuracy between upper and lower case forms of single words. An even more dramatic problem for the envelope representations comes from the finding that destroying the envelopes by alternating case does not affect how long it takes to report or read a word (Smith 1969; McClelland 1976; Adams 1979). These results provide evidence for the view that either it is the recognition of individual letters which is important rather than global shape, or that competitors are more abstract than a strict envelope definition allows.

One point on which proponents of both envelope and individual letter similarity explanations agree is the importance of the onset letter. In forced choice judgments of target/probe similarity, although it is impossible to distinguish envelope effects from those of letter similarity, it is clear that target and probe must share onset letters for priming to be observed (Marchbanks and Levin 1965; Rayner 1976).

Evidence for the existence of the envelope as a level of representation relevant to the processing of words appears to be at best unclear. It may be useful to consider that thirty percent of the 2065 four letter words of English can be accounted for by only two envelopes, those of the words, *neon* and *bone*. For short words, envelope information defined solely in terms of ascenders and descenders, in the absence of specific letter details, does not substantially limit the potential candidate set.

On the other hand the fact that readers can estimate, with reasonable accuracy the length of words presented parafoveally (Rayner 1975) does suggest some level of early global consideration. The white space which surrounds the beginnings and ends of words makes them more visually distinct than the intervening characters. Furthermore, the beginnings and ends of words contain very useful information. Word onsets provide the start of pronunciation and word endings often carry syntactic information. Though this latter source of information is often quite redundant in context, in isolated recognition it may be more important. Outer letter frames may therefore be a relatively better contributor to perception than underdefined complete word envelopes.

2.1.2 Single letters

"We cannot have a position free awareness of the identity of a letter

> in a word anymore that we can experience the redness of a book cover in isolation, seeing involves beliefs about stimulus objects." Henderson (1982).

In line with Hendersons' comment competition effects concerning single letters appear to operate in relation to the letters context. Besner (1984) has shown that same/different judgments based purely on physical criteria are influenced by whole string familiarity: *FBI/fbi* takes longer to respond to than *FDM/fdm*. Subjects appear unable to process the three letters *F*, *B*, *I*, without some interference from the fact that holistically the uppercase letters are a meaningful acronym. Interference from semantics makes *FBI* and *fbi* dissimilar because only *FBI* is perceived as a meaningful acronym. The three-letter pairs that do not represent acronyms can be compared purely as letter representations. Despite the fact that the real world knowledge concerning the letters FBI is antagonistic to the letter comparison task subjects are unable to suppress it. This finding demonstrates our sensitivity to context even when such sensitivity is orthogonal to the task.

For most tasks however, an awareness of the context in which single letters occur is beneficial to processing. Single letter probe recognition studies have demonstrated that for words and pseudo-words, the probability of correctly reporting all the letters is greater than one would expect if all the letters were processed independently (McClelland and Johnston 1977). There appears to be some single letter positional redundancy operating within the recognition system. This redundancy prevents us from having to adjudicate evenly between twenty-six possible letters at each letter position. Letter report accuracy is more accurate when letters occupy positions in which they commonly occur (McClelland and Johnston 1977). Strings with higher positional redundancy are named faster (Masson 1978), classified in LDT faster (Henderson and Chard 1980), and identified tachistoscopically more accurately (Masson 1978; Massaro, Venexky and Taylor 1979) than less redundant strings. The speed of letter search through nonwords is also affected by positional letter redundancy (Masson 1975; Massaro, Venexky and Taylor 1979). In terms of the competition metaphor, we may consider that the strength of a single letter competitor depends on the frequency with which it occurs in a specific letter slot.

2.1.3 Letter clusters

Suggestions as to the nature of larger than letter processing units range from bigrams (Segui and Grainger 1991), which consist of two adjacent letters, up to orthographic neighbours (Coltheart 1978) which are only one letter different from a whole target word (*e.g. bed, wed, led, fed, red, bad, bud, bid, beg, bet*). Representations as large as neighbours we will leave until the next section. The remaining sub-word representations can be classified as appealing to either sequential redundancy or pronunciation.

Bigrams are the most frequently used measure of sequential redundancy (Henderson 1982). They are also the smallest unit. A frequent bigram should be easier to recognise than a rarer one. In tachistoscopic report (McClelland and Johnston 1977), LDT and word naming (Andrews 1992), however, bigram frequency is found to have no effect on response time or accuracy.

A number of researchers have considered the role of syllables in visual lexical access (e.g. Forster and Chambers 1973; Fredrikson and Kroll 1976; Spoehr and Smith 1973, 1975). If words were parsed into and processed via syllables, then words with more syllables should produce longer response times, but this is not found (Massaro 1975). The absence of syllable number effect on response times does not however preclude syllable analysis in lexical access. Taft (1991) suggests that the emphasis is on first syllables as the access representation or code. Polysyllabic words could then have an advantage over monosyllabic words of the same letter length since their access codes would be shorter (e.g. *tiger* has *ti* and *ger* as its access codes but *train* will have only the larger code of *train*). Evidence for the use of first syllables of a stem word as access codes comes from a variety of experiments. For example nonwords which are the first syllables of monomorphemic words (e.g. *ath* from *athlete*) take longer to reject than nonwords which are not the first syllabic part of a word (e.g. *aw* from *awful*) (Taft and Forster 1976). We may consider that syllable nonwords experience competition from the real word they are taken from, whereas non syllable nonwords have no lexical competitors.

Amongst linguists there is much controversy surrounding the phonological definition of a syllable (see Taft 1991). This controversy was one of the motivations behind the specification of the BOSS (Basic Orthographic Syllable Structure) hypothesis, put forward by Taft (1979) and supported by Prinzmetal, Hoffman and Vest (1991). The BOSS hypothesis states that printed words are encoded via orthographic syllabification and that whole words are represented in the lexicon as syllables constructed within the constraints of English orthography. The BOSS of a word is the first syllable encountered by left to right letter analysis after the stripping of any prefixes or pseudo-prefixes, plus as many consonants, following the first vowel of the word, as orthotactic factors will allow without disrupting the morphological structure of that word (Taft 1979 pp. 24). For example the BOSS of *demon* is *mon* because *de* is stripped as a pseudo-prefix, and *lant* is the BOSS of *lantern* because *t* can be added to the first syllable without violation of English orthography, or morphology.

All stimuli sharing a BOSS are assumed to share the same access code (e.g. *deMON*, *MONarch*), we may therefore expect competition to occur between a candidate set of BOSS sharers. Jordan (1986) demonstrates however that *lemon* which has the BOSS *lem* primes *monarch* just as well as *demon* does. If a BOSS was the important level of representation *demon* should prime *monarch* but *lemon* should not. It appears therefore that although the BOSS of a word may not always be important to recognition multi letter clusters such as *mon* operate in a non-position specific manner.

2.1.4 Whole words

If words compete at a whole word level of representation then it is logical to suppose that the active candidate set for the recognition of a visually presented word will contain visually similar words: those having greatest orthographic overlap with the target. Coltheart, Davelaar and Besner (1977) put forward the N metric as a measure of orthographic similarity. N corresponds to the number of new words of equal letter length generated by replacing any letter of a target word with another letter in the same position (e.g.*bird* has the neighbours,*gird, bard, bind, biro*). N indicates the number of orthographic neighbours a word has and thus represents the word's neighbourhood density. Neighbourhood size is assumed to be a lexical construct. The measure was devised by Coltheart et al. to study the role of lexical level competition.

At a whole word level of representation there can only be one winner of the competition process for accurate recognition to occur. Sharing visual overlap with other words, as neighbours do, will make a target less distinct and thus less separate from members of its candidate set. We may therefore expect words from large neighbourhoods to experience stronger competition than words from smaller neighbourhooods.

Whatever Coltheart et al. intended, neighbours have elements in common, specifically the sublexical components from which words are formed. Words from large neighbourhoods of lexical competitors are, at the same time, those words made up of common bigrams, trigrams and other sublexical units of representation. The evidence presented in the preceding sections of this Chapter suggests that commonly occurring sublexical components show effects of orthographic redundancy. The sublexical character of neighbours therefore predicts ready recognition of large neighbourhood words, a complete contrast to the competition predicted by their lexical level character. The literature contains evidence to support both lexical and sublexical predictions independently, but the two cannot always be reconciled.

Coltheart (1977) reported that neighbourhood density had no effect on word responses in LDT but did have an interfering effect on nonword rejection. Andrews (1989) has shown however that adding the manipulation of word frequency to the studies of Coltheart results in a neighbourhood size effect for words as well as nonwords. Low frequency words with many neighbours were recognised as quickly as high frequency words. There was still no effect of neighbourhood size for high frequency words. The facilitatory effect was observed when both neighbourhood size and word frequency were manipulated. Although Andrews reports neighbourhood effects for words, this can not be taken as evidence for a lexical level influence since neighbours may compete lexically but conspire sublexically. It could be this sublexical similarity (as Segui and Grainger (1991), suggested) that produces Andrews' neighbourhood effect. Andrews (1992) has shown, however, that bigram frequencies do not affect the LDT and naming responses which are influenced by neighbourhood size. The implication from Andrews work is that units as large as neighbours may have some importance to the word recognition system additionally or independently of their sublexical components.

In contrast to the findings of Andrews, neighbourhood size has been shown to inhibit visual word recognition in form priming (Colombo 1986; Segui and Grainger 1991; Forster, Davis, Schoknecht and Carter 1987; Forster and Davis 1991). Furthermore, LDT responses and error rates have been found to increase as soon as a target word has at least one higher frequency orthographic neighbour (Grainger, O'Regan, Jacobs and Segui 1989; Grainger 1990; Grainger and Segui 1990).

Snodgrass and Mintzer (1993) found both facilitatory and inhibitory effects of neighbourhood size for the same visually fragmented targets depending on the method of stimulus presentation they employed. In an ascending method of limits task, subjects are presented with gradually less fragmented images and have to guess at the word identity after each stimulus presentation until they are correct. Neighbourhood size was found to be facilitatory for this task, large neighbourhood words were correctly identified on more fragmented image levels than were low frequency words. When fixed level fragments were presented, neighbourhood size was inhibitory: large neighbourhood stimuli were identified at much more complete image levels than small neighbourhood words. Snodgrass and Mintzer present an explanation for the opposing neighbourhood effects in terms of the hypotheses subjects entertain concerning a word. The ascending method of limits gave subjects the opportunity to test and reject early hypotheses of strong neighbourhood competitors. The fixed level fragment presentation did not provide any feedback, or opportunity for people to disconfirm their initial hypotheses.

The implication from Snodgrass and Mintzers work is that in experimental paradigms using inferior quality stimuli, neighbourhood size is inhibitory because larger neighbourhoods contain more frequent competitors (Bard 1990) and it is these strongest competitors that cannot be rejected. When normal processing is able to run to completion strong non-target candidates can be rejected. It is still unclear however why the competitor set would then become facilitatory. Furthermore, though Snodgrass and Mintzer's explanation may be relevant to the inhibition of form priming, it does not explain Grainger et al's. lexical decision findings. In these studies stimuli are not degraded, but neighbours apparently compete with targets.

2.1.5 Summary: competition effects

The lexical access literature appears to contain much evidence for the existence of competition between words at a variety of representational levels. Two factors which are important in lexical access are frequency and predictability. Just as high frequency words are recognised more quickly than less frequent ones, so more frequent sublexical components can be recognised more quickly than less frequent ones. Similarly letters with a high context conditioned probability appear easier to recognise than letters in less probable or random contexts. The more frequently occurring or context-predictable a letter pattern is, the more words it will appear in, and so the larger its competitor set will be. If such letter patterns are easier to recognise than less probable ones, we may infer that the existence of a large competitor set can be facilitatory, that is, at least sublexically candidates conspire.

Word level non-target candidates are reported to be both inhibitory and facilitatory to target recognition. The overall picture from the literature concerning lexical candidate sets is therefore very unclear. There are many questions which remain unanswered and effects that are without satisfactory explanation.

The conspiracy effects of lexical neighbours found by Andrews (1989, 1992) are found only for low frequency words and not for high frequency words. Andrews suggests that since high frequency words are so easily recognised, they have no need to benefit from neighbourhood activation. That for high frequency words, there is effectively a ceiling effect. This explanation is quite vague and underspecified, it offers a description without a satisfactory mechanism. In this thesis we attempt to demonstrate and explain how and why competition or conspiracy effects differ for high and low frequency words.

The distinction of different processing levels of representation is a legacy from early models of lexical processing. Modern approaches, such as those motivating connectionist models give greater consideration to the stochastic nature of lexical experience. Later in this thesis we give fuller consideration to the intercorrelations of lexical variables. For now we will just consider the confounding of neighbourhood size and bigram frequencies. Andrews (1992) claims that neighbourhood size conspiracy operates independently from sublexical effects and is the demonstration of true lexical level conspiracy. It is difficult to see how words

	Word recognition task		
Process	LDT	Immediate naming	Delayed naming
Access	Yes	Yes	No
Decision	Yes	No	No
Lexical pronunciation	No	Yes	Yes
Sub-lexical pronunciation	No	Yes	No

Table 2–1: Locus of effect as determined by task comparison.

could conspire if not though their shared constituents. In this thesis we address this issue in relation to another problem, the problem concerning the reversal of neighbourhood effect from inhibition to facilitation.

The neighbourhood size conspiracy effect found by Andrews is reportedly weaker in word naming than in LDT. Effects which are found to be present only for the decision task are often attacked as biasing the decision process rather than affecting lexical access. In this thesis we therefore consider what it is that neighbourhood size influences, and why the influence is greater for LDT than naming.

2.2 Lexical and sublexical effects

One of the ways in which models of word recognition differ is by the number of processing routines they include. As described in Chapter 1, the original dual route model (e.g. Coltheart 1977) has a visual process which directly accesses the mental lexicon, and a phonological process where grapheme-phoneme correspondences are used to assemble pronunciation. In contrast, Seidenberg and

McClelland's (1989) connectionist model and Glushko's (1979) analogy model both have a string of letters pronounced with reference to similar lexical entries, and so they require no non-lexical routine.

In order to show that more than one way of processing words exists, it is necessary to find evidence for factors which influence any one method and not others. Models often appeal to between-task comparisons for structural evidence and information. If two tasks are presumed to have a common process which a third task does not share, then an effect manifest in the first two tasks and not the third may be attributed to this process.

2.2.1 LDT and Naming

LDT and naming tasks are often used in combination to examine lexical access. The logic behind their joint employment is that the process they have in common is lexical access. Manipulation of a variable which affects both tasks, therefore, may be affecting the lexical access process, if by lexical access we mean the mapping of a string in its entirety onto a representation in lexical memory. LDT demands lexical access because asking whether a string of letters is a word or not is equivalent to asking whether such a string exists in lexical memory. In addition to access, LDT also requires a decision process to enable a response to be made. This decision process is not common to naming, so any variable having an effect on LDT and not naming is likely to be affecting the decision process. Whether naming responses are requested immediately or after a delay, naming involves the generation and execution of a pronunciation. A variable which affects overt pronunciation will show an effect on naming responses but not on LDT. Since it is possible to pronounce nonwords, immediate word pronunciation could be performed at a sub-lexical level by consulting grapheme-phoneme correspondences or by analogy to word bodies or rimes. In a delayed naming task subjects are made to withhold responses until a given cue. There is therefore more time for complete lexical access than in immediate naming. At least for real word pronunciations, variables which affect sub-lexical pronunciation processes are more likely to affect only immediate naming. Delayed naming allows access processes to run to completion and so we do not expect access effects to influence delayed naming responses. Comparing results from LDT and naming, therefore, provides some indication as to the locus of an effect; access, decision, lexical or sub-lexical pronunciation (see table 2–1).

Conclusions as to the locus of effects in lexical processing are not, however, as straightforward as the task comparison approach suggests. The processing stages are most likely not independent modules and the processes used to perform a task may vary within and between subjects. In naming words, subjects may use lexical and sub-lexical pronunciation or both rather than just one or the other. In LDT subjects may use the existence of pronunciation codes in memory as the basis for a positive "WORD" response.

2.2.2 Lexical and sub-lexical naming

Lexical involvement in naming responses is implied by word frequency effects in naming. The word frequency effect is generally believed to arise from the lexicon and so is indicative of lexical access. Frequency effects for word naming are however weaker than those observed in LDT (Frederiksen and Kroll 1976). The weaker effect may be due to the use of both lexical and sub-lexical processes.

Although naming a list of real words may be performed lexically or sub-lexically, naming a list of nonwords is more likely to use a sub-lexical process since none of the stimuli will have lexical representations. If nonwords are named lexically it must be by analogy to those lexical items which they most resemble (Glushko 1979). Pseudohomophones are nonwords which will have whole word phonological representations within the lexicon. These phonological representations will have the lexical characteristics of their real word base. Naming times for blocks of pseudohomophones show no effects of base word frequency (Taft and Russell 1992). Naming times for mixed lists of words and pseudohomophones do reveal an effect of base word frequency (McRae, Jared and Seidenberg 1990). The list composition differences between the Taft and the McRae studies appear to have influenced which naming process subjects adopt. In the situation where all the stimuli were nonwords, a sub-lexical process was favoured and so no word frequency effects emerged. When real words are added to the stimulus list, then a lexical process is employed to some extent and a word frequency effect is found. Moreover a lexical process is believed to take longer to complete for new or nonwords than a sub-lexical process. The slow subjects of the Taft and Russell study are therefore the ones most likely to be using lexical processing, and thus more likely to exhibit a word frequency effect. This is what is found, the slow subject responses of Taft and Russell show clear word frequency effects comparable to those of McRae et. al.

2.2.3 Pronunciation influences in LDT

Irregularities in spelling-to-sound rules can have a detrimental effect on pronunciation: the more irregular a word is the slower a response will be made (Parkin and Underwood 1983). Coltheart, Besner, Jonasson and Davelaar (1979) failed to obtain phonological effects of irregularity in LDT and concluded that LDT is mediated by direct visual access and not by phonology. Parkin (1982) has since demonstrated that producing the phonological effect in LDT is contingent on the definition of exception words. The correspondence of *i* to */ai/* is an exception, but the correspondence of *ign* to */ain/* is not. Common exception words such as *sign* behave as regular words in LDT. Phonological recoding only appears to be active in LDT for very unusual spelling to sound matches (Parkin 1982).

Exception words are often not just spelling-to-sound irregular but are also graphemically unusual, e.g.*aisle*. The odd visual characteristics of irregular words could just as easily be what marks them out from the more familiar letter sequences of regular words. If this were the case, then evidence for phonological mediation in LDT would be undermined. Orthographically regular, phonologically irregular words e.g. *pint*, do, however, take longer to respond to than purely regular words (Parkin and Underwood 1983). Irregular words would appear therefore to gain little advantage from having a regular orthography. Perhaps, orthography is not at issue and processes underlying LDT can regularly involve phonological mediation.

Another source of evidence for phonological processing in LDT is the pseudohomophone effect. Pseudohomophone foils such as *nale* take longer to reject in LDT than non-homophonic controls (Meye and Gutschera 1975). The string *nale* appears to activate some phonological representation which increases its similarity to *nail*, or which just registers "WORD" in some more abstract way, and thus makes it harder to class as a nonword. The apparent phonological effect occurs, however, on "NO" trials which are always very slow and so may reveal little about phonological mediation in normal word recognition. A categorisation task enabled Van Orden (1987) to observe phonological effects in "YES" responses. Significantly more false positives were made to homophones such as *rows* (as an exemplar of *flower* than to spelling controls such as *robs*. Homophones such as *meet* which is similarly spelled to *meat* produced more false positives than homophones like *rows* which do not share much orthographic overlap with their partners (*rose*). It seems therefore that both orthographic and phonological representations can influence decision responses.

2.2.4 Models of word naming

The picture of word identification which emerges from the experimental evidence is much less clear cut than any assumed functional architecture. To accommodate findings, models have become very broadly defined with the result that it is increasingly difficult to generate predictions which are uniquely tied to any one theoretical account (Brown and Besner 1987; Norris 1985; Henderson 1985; Seidenberg 1985). As evidence accumulates which contradicts a model's assumptions, so the model mutates or relaxes to absorb the new information. Probably the best example of the increasing proximity between once divergent views concerns the issue of single or multiple pronunciation routines.

To account for nonword or new word pronunciation, the assembled phonology process of a dual route model traditionally employed very small translation units. Single letters or letter clusters were mapped onto a single phoneme. In contrast, the process of lexical analogy originally appealed to much larger translation units favouring vowel and consonant clusters.

Ambiguous nonwords (e.g. FOTH) can have more than one pronunciation pattern (c.f. MOTH vs BOTH). Experiments using such strings have demonstrated that it is possible for the string to be given the irregular pronunciation (Kay and Marcel 1981; Glushko 1979). This evidence for the pronunciation systems' sensitivity to word bodies (Patterson and Morton 1985) or to word rimes (Treiman and Chafetz 1987) has led to the inclusion of these larger representations in revised dual route accounts. Another important implication from the work of Glushko was the demonstration that the pronunciation of a word is influenced by knowledge of other similarly spelled words. Performance was determined by the consistency of a spelling-to-sound correspondence defined in terms of a neighbourhood of similarly spelled words and not by whether the correspondence was governed by a specific rule. The dual route response was to assume that multiple rules could be associated with a given orthographic unit and applied on a probability basis (Patterson and Morton 1985).

Analogy routines have similarly had to shift position to place greater emphasis on smaller units such as phonemes. This is necessary to cope with findings such as nonwords like *yook* with a predominantly irregular body level analogy (e.g. *book, look, hook, cook, rook*) being most frequently assigned a regular pronunciation (e.g. *yuke*). For a single route model to accommodate this finding *yook* must be pronounced not by analogy to its body sharers but by analogy to other words containing the *oo* bigram, such as *hoop, loot.* Readers are also easily able to assign pronunciation to nonwords which have no body sharers at all (e.g. *pelb* might be pronounced by analogy with *pelt* and *elbow*). The idea that orthographic and phonological lexical representations are segmentable, or are stored as segments is an extendable principle. Small units (even graphemes and phonemes) were therefore always available within analogy accounts but the original emphasis was on units as large as possible up to a word level.

It would no longer seem appropriate to characterise the relationship between analogy and rule-based accounts of phonological translation as one of opposition. The blurring of the distinction between frameworks makes it more difficult to accept any one model as a truer account than any other. The areas of greatest similarity between accounts are those we may feel most confident in accepting as a likely reflection of events. The convergence of approaches illustrates that the word recognition system must deal with various levels of representation (e.g. the Shortlist model by Norris 1994; and the model of Plaut and McClelland 1993). Convergence also indicates that the processing of strings is influenced by knowledge of other words. This knowledge may be a direct influence of other lexical entries or an indirect influence via the probabilities of rules. The interactive activation model of McClelland and Rumelhart (1981) and the subsequent connectionist models which have followed it (e.g. Seidenberg and McClelland 1989) also include in some fashion the principles of multiple-level representations and of lexical experience.

2.2.5 Summary: lexical and sublexical effects

Lexical decision and naming tasks are often used in combination. The logic behind their joint employment is that naming involves phonological processes and

LDT does not. If a variable affects naming and not LDT, it may be assigned a role in phonological encoding. This sort of task comparison paradigm has been used to support box and arrow type models of lexical access. These models modularise various components of word recognition, such as phonological encoding. Experimental evidence has shown, however, how both LDT and naming can be potentially affected by lexical or phonological variables. Models have, therefore, had to become less rigid to accommodate the new knowledge. A result of this relaxation is that it has become increasingly difficult to distinguish between the models in terms of what they predict.

In this thesis we examine competition effects across a variety of tasks and consider how well models accommodate the data obtained. In keeping with the present theoretical trend, we appeal to connectionist architectures. The specific issue with which we are concerned is the effectiveness of "single route" multilayer connectionist models (e.g. Seidenberg and McClelland 1989; Plaut and McClelland 1993) in dealing with between task differences in competition effects.

2.3 Strategies in visual word recognition

In the previous section we mentioned that the word/nonword content of a stimulus list affects the emergence of a frequency effect for naming pseudohomophones (Taft and Russell 1992; Jared, McRae and Seidenberg, 1990). Response instructions give the subject explicit information concerning how to go about the task. The nature of a stimulus list may provide an implicit source of information.

2.3.1 Strategies

In learning and problem solving, people rarely behave in a random or accidental fashion but plan and execute their responses with some purpose. People generally prefer to operate in the least effortful and most efficient manner possible. When asked to introspect on how a new problem is solved, subjects are often embarrassed or ashamed of their initially clumsy solutions (Wood 1978). If subjects feel a more direct route to the goal is possible, then they seek to find it. Once an individual has a method for approaching a solution we may describe the behaviour as a strategy.

Explicit instructions given to subjects may give some clue as to how to solve the presented problem. These instructions will therefore influence the strategy people adopt, and thus the outcome they produce. Dagenbach, Carr and Barnhardt (1990) found that newly learned words succeeded as semantic primes, only, if subjects were told to use them to predict the next word. The same new words were completely ineffective as primes when subjects were told only to bring the word to mind. Subjects showed semantic priming effects only when they used a strategy that utilised their new knowledge to make a prediction.

Explicitly stated time constraints can also affect the strategies subjects will use to solve a task. Subjects placed under time pressure may make less detailed assessments of a problem, or may be able to complete only a portion of the necessary processing. The effects of masked repetition and semantic priming (Carr and Dagenbach 1990), and the effects of orthographic neighbourhood size (Snodgrass and Mintzer 1993) have been demonstrated to be both inhibitory and facilitatory. The nature of these effects depends on the processing time constraints imposed on subjects.

At one level of description a strategy is a "programmatic assembly of operations aimed at a single goal" (Woods 1978). This type of definition reasonably describes the intentional strategies described above (e.g. using a prime or speeding responses as directed) which are based on explicit pre-task information. The information acquired is assimilated into any solution strategy attempted. The definition of strategy implies however the concept of choice or optionality. This makes it difficult to relate strategies to implicitly acquired information.

Yet the implicit learning literature provides many demonstrations of subjects' ability to pick up on regularities or patterns within materials without having an explicit awareness of what those patterns or regularities are (Berry and Broadbent 1984, 1988). People have been shown to learn novel grammars and develop accurate expectations as to what is legal and illegal for a particular grammar without explicitly knowing any rules (Reber 1967, 1976; Reber and Allen 1978; Reber, Kassin, Lewis and Cantor 1980). We would not wish to suggest that in implicit learning subjects make any conscious choices about their behaviour or responses. Rather we would suggest that implicit learning is a demonstration that intellectual activity is extremely adaptable to the local environment. The "rules" are not fixed by common or pre-task experience, but take their form and procedures from rather more specific experience, such as the materials of a particular experiment. Intellectual operations such as strategies may therefore arise as a natural or automatic product of experience.

2.3.2 Stimulus lists

Most investigations of visual word recognition use isolated word presentation. For these studies the context, or local environment, of any presented word will be the other words used in the study, the items which make up the stimulus list. Stimulus lists naturally vary between studies. It is possible that much of the variation within the reported lexical access effects may be the result of subjects' sensitivity to factors within particular stimulus lists. In visual word recognition tasks, subjects may adopt local response strategies. If this is so, then it should be possible to demonstrate different response times to the same target word by varying the stimulus list context in which it appears.

Experiments have certainly shown that the effect of one word's frequency depends on the frequency of other stimulus list items. Reaction times for low frequency words are the same in mixed lists and in pure low frequency lists; in contrast, high frequency words are responded to much faster in pure lists than in mixed lists (Gorden 1983; Dorfman and Glanzer 1988). The suggested explanations for a relative effect of word frequency centre on the setting of evidence criteria. High frequency words are assumed to have a higher resting activation level (McClelland and Rumelhart 1981; Rumelhart and McClelland 1982) or lower activation threshold (Morton 1969, 1970) than low frequency words. High frequency words require less evidence, that is, less information about the stimulus is required to exceed threshold or activate a word, so a low evidence criterion can be set. Low frequency words require a much higher criterion since greater amounts of evidence are required to fire or activate low frequency words. In mixed frequency stimulus lists a low criterion would seriously compromise subjects' accuracy for low frequency words. The literature suggests that in viewing a mixed list, subjects have a high criterion, so that low frequency words can be responded to accurately but high frequency words "suffer".

In a lexical decision task the problem is to discriminate words from strings which are not words. Any word dimension which promotes the word/nonword difference will be useful to the discrimination task. The manner in which stimuli are processed and the way in which the results of processing are translated to word/nonword responses may vary greatly according to the content of the stimulus list. For example if all the nonwords in a stimulus list were illegal strings and all the words were legal, then the decision could be made on the basis of how common the letter patterns were. If nonwords are very wordlike then a strategy of discrimination based on letter pattern commonness would be less effective. Stone and Van-Orden (1993) have shown that increasing the wordlikeness of nonwords increases the word frequency effect and decreases errors to oddly spelt words. It appears possible for subjects to adjust their decision criteria in response to the experimental conditions with which they are faced.

The structure of stimulus lists used in lexical decision has also been shown to influence the use of phonological information. Inconsistent findings concerning the pseudohomophone effect have been traced by Besner et al. (1985) to the variability in types of word/nonword stimuli employed in different experiments. If the stimulus conditions encourage subjects to rely on phonological information in making the decision then a pseudohomophone effect is found. If the stimulus conditions do not encourage phonological processing then a pseudohomophone effect is not found. Waters and Seidenberg (1989) show that the use of phonological information in making lexical decisions depends on the type of words included in the test list. The inclusion of unusually spelt words encourages subjects to use a phonological strategy. Stimulus lists which exclude unusually spelt words are shown to encourage an orthographic strategy. Seidenberg (1989) even raises the possibility of strategic behaviour being responsible for the stem morpheme effects reported by Taft and Forster (1976). Taft and Forster show that nonwords containing real word stem morphemes (e.g. *delease*) are classed as nonwords more slowly than nonwords which contain only non-stem fragments. Taft and Forster put this forward as evidence that a stem is stripped of its prefix in processing. Seidenberg suggests that subjects may decompose the stimuli only because doing so facilitates the word/nonword decision.

Strategic behaviour is often only acknowledged *post hoc*. When unexpected results arise or predicted ones do not, then strategies are turned to for an explanation. A theory that too freely invokes strategies can however account for any possible empirical pattern (Besner 1984). Strategic explanations have therefore gained a bad reputation because they are used to patch up the inconsistencies between prediction and result and thus undermine the value of the empirical test.

It may be equally dangerous to ignore strategies. If we refuse to consider the type of strategy subjects would be likely to adopt or the extent to which different attributes of the materials would invite strategic treatment, we will be unable to distinguish stimulus and strategy driven effects. Many complexities and contradictions within the literature which we standardly consider to be stimulus driven may be unraveled by the serious consideration of strategies. Very little information is usually given concerning the specific stimuli used in experiments, but the evidence discussed in this thesis suggests such information may be vital to understanding subjects' strategic behaviour in task performance.

2.3.3 Summary: strategies in visual word recognition

In problem solving tasks, people generally form some sort of response strategy based on the explicit information they are given, and on their initial experience of the problem. If the tasks presented in word recognition experiments are considered as problems then subjects may develop response strategies to aid their performance. These strategies will be formed from the explicit instructions subjects are given about the task, and from the implicit information they gain during the course of the experiment.

If we are to address the issue of stimulus list influences, we need to establish which sorts of stimulus dimensions are influential, and how the mechanism of influence operates. In performing a word recognition task, subjects have two sources of information. Firstly they have knowledge about words, their lexical memory. Secondly they have information obtained during the experiment, a more local type of information. The approach we take in this thesis to investigate stimulus list effects is to consider the disparity between lexical and local experience.

Chapter 3

General Methodology

The research presented in this thesis is experimental and involves the measurement of responses (usually reaction times) on tasks believed to involve lexical access. Seven of the eight experiments use isolated word presentations, while the eighth has words presented in sentence contexts. The general methodology of the thesis work is to construct very tightly controlled stimulus lists which allow a single variable of interest to be manipulated while as many other known variables as possible are held constant. In this chapter we will firstly describe how and why we made the stimulus lists used in this thesis. We will then give details of the specific versions of tasks these stimulus lists were used for.

3.1 Stimulus specifications

One of the most important questions we can ask about visual word recognition is, what it is about a particular word which makes it easier, or more difficult to recognise, than the next word? What are the factors which determine how we recognise visually presented words? The way we investigate this problem is to firstly construct a hypothesis about a variable we believe to be relevant, and then we attempt to determine how this variable influences recognition. By constructing test materials which differ on the dimension of the investigated variable, we attempt to isolate the influence of this variable in word recognition. The ideal materials for word recognition tasks would be words which differed from one another only on the dimension being manipulated. Since this is not possible, and characteristics co-vary, we find sets of words which are as similar as possible in many different ways but that show distinct differences on the dimension of the manipulated variable. In the following sections we describe six of the word dimensions that are manipulated or controlled in the thesis experiments. The six variables are all relevant to more than one of the experiments. Four other variables (word frame, Near-non-neighbour, word shape and sentence context), are specific to the particular experiment in which they are examined; for this reason these latter variables are described fully in the relevant experimental chapters.

3.1.1 Word length

The simplest and perhaps most obvious dimension on which to equate words is length. By length we can mean number of letters or number of syllables. All of the stimulus items used in the experiments are four letter words, and as far as possible they are monosyllabic.

3.1.2 Word frequency

Frequent words are generally considered to be processed more quickly or more easily than rare words. It is therefore important to control for word frequency when investigating the potential effects of other variables. In experiments 4, 5, 6, and 7, of this thesis word frequency is manipulated (high/low), in all four other experiments word frequency is carefully controlled.

The frequency of usage of a word in the language is determined by frequency counts in large corpora of texts. The published corpora counts, such as those of

Kučera and Francis (1967), provide an operational definition of frequency. Since the sources of these counts will not correspond exactly to everyone's reading experience, no one would suppose any one count is the truth. Rather, these counts simply provide good estimates of how often an experimental subject will have encountered different words. In the experiments presented here word frequency was determined by the Kučera and Francis (1967) number of samples norm. The number of samples norm is generally considered to be the best of the printed word frequency measures. It indicates a breadth of usage, and generally correlates with other behavioural data more strongly than any of the other frequency variants.

3.1.3 Orthography

Neighbourhood size

Words can differ in how common their spelling pattern is. For example the sequence *ill* occurs in words such as, *bill, dill, fill, hill, kill, mill, pill, sill, till, will, chill, shrill,* etc. But a sequence such as *eon* occurs in far fewer words. One measure which captures this orthographic frequency is neighbourhood size, Coltheart's N metric (Coltheart, Davelaar and Besner, 1977). N corresponds to the number of new words generated by replacing any letter of a target with another letter in the same position (e.g. *word: cord, wood, worm* etc.). In experiments 3, 4, 5, and 8, we investigate the manipulation of orthographic neighbourhood size (large versus small). In all four other experiments neighbourhood size is controlled across conditions.

Bigram frequency

A second measure of how frequently letters occur together in words is bigram frequency. A letter bigram is a pair of letters or letter and space in written

Attributes	Correlation
Imagery $\times$ concreteness	0.83
Imagery $\times$ meaningfulness	0.72
Concreteness $\times$ meaningfulness	0.56

Table 3–1: Inter-correlations of the attributes; imagery, concreteness, and meaningfulness, as found by Paivio, Yuille and Madigan (1968).

English. The frequency of a bigram is determined by finding the number of times a particular letter pair occurs in printed text corpora. The frequency of a bigram is therefore influenced by the frequency of words in which it occurs. Word bigram frequency is usually calculated by finding the mean frequency of the individual letter bigrams within each word.

The letter bigrams used to determine the word bigram frequency scores in this thesis were taken from a million word sample of the USENET news group. The nature of this corpus means that some words will be more common here than in everyday language. Yet any corpus will have idiosyncrasies. For example a corpus of 10,000,000 words from the *Wall Street Journal* was found to have "million" as its tenth most common word. Like the word frequency counts the bigram frequency counts are estimates.

It has been suggested that bigram frequency may account for some of the effects attributed to neighbourhood size, since words with large neighbourhoods are those with a common orthographic pattern and thus common bigrams. In experiments 4, 5 and 7 we attempted to isolate the effects of neighbourhood size from the influence of bigram frequency by holding bigram frequency constant across conditions while varying neighbourhood size (small versus large).

3.1.4 Concreteness

Words which refer to persons, places and things that can be seen, felt, smelled or tasted are usually very meaningful and concrete. Words such as, *democracy* and *independence*, refer to more abstract concepts which cannot be experienced by our senses. Measures of concreteness or meaningfulness are determined by getting subjects to judge the words on a seven point scale e.g. Paivio, Yuille and Madigan (1968). Several different measures, *concreteness, imagery*,and *meaningfulness*, are often considered to measure the same underlying variable. Paivio, Yuille and Madigan (1968) measured 925 words on all three measures. As Table 3–1, shows a high correlation is found between the measures.

Words that attain different scores on the above mentioned attribute scales, and that differ between scales by more than two scale points, fall into one of the following three categories;

- Those having strong emotional or evaluative connotations (eg. agony, grief, passion).
- Those referring to fictitious characters (eg. devil, ghost, phantom).
- Very unusual words (eg. kine, gadfly, surfax).

Since none of the words in the present material sets fall into any of these categories, it seems reasonable to employ just one of these measures. The measure of "Concreteness" was chosen because Paivio, Yuille and Madigan found it to be the most reliable, and the most obviously bipolar of the three measures. Because the present materials include many words not covered by Pavio's study, it was necessary to replicate the concreteness survey for our stimulus lists.

	Number of items	Correlation	P
Paivio	36	0.65	0.00
Colorado	51	0.06	0.80
Concreteness	78	0.91	0.00
Imagery	78	0.80	0.00

Table 3–2: Correlations between the concreteness scores of the present study and measures reported in the MRC database

The concreteness survey

The procedure for determining measures of concreteness for the present materials was the same as that used by Paivio Yuille and Madigan: a modified version of that used by Spreen and Schulz (1966).

Instructions and items were presented as a booklet to 15 skilled readers. Each page contained 20 words in uppercase letters on the left hand side of the page. The words were listed in five different random orders. Opposite each word was a seven point bipolar scale with the extremes labeled "Highly Abstract" and "Highly Concrete". The booklet contained one hundred words to be rated; the sixty-four words used for the eyetracking experiment (see Chapter 5) and thirty-six words which had relatively abstract ratings on the Paivio norms. The latter thirty-six words were included because the eyetracking materials consisted of mostly concrete nouns. The inclusion of Paivio's words allowed tests of reliability between Paivio's results and those of the present study. The instructions were stated on the cover sheet of the booklet.

The concreteness for each word was determined by finding the mean rat-

	Voiced onset sound		
	Stops	Fricatives	Approximants
Andrews 1989	586		576
Andrews 1992	517		519
Chumbley & Balota 1984	556	562	518
Waters & Seidenberg 1986	515	616	496
Means	535	589	506

Table 3–3: Mean naming times (ms) for items with voiced onsets across six different studies

ing and multiplying it by 100. The original Paivio norms, established by the same method used for the present study, correlate significantly with the present ratings, (for 36 items $r = 0.65$, $p < 0.01$).

3.1.5 Word onset

Naming latencies are measured by the triggering of a voice key at word onset. Named words can differ in inherent latency because of the amount of energy produced at their onsets. The greatest contrast is between a voiced non-stop and a voiceless stop consonant sound. A word with a voiced onset has energy released at onset where as, for a voiceless stop no energy at all is released until the vowel sound. Voiced stops release voicing earlier than voiceless ones, while fricatives release energy throughout their articulation. For these reasons, onset sounds should be carefully controlled so that apparent naming effects are not just an artifact of the uneven distribution of onset types.

Within the reading literature, surprisingly little concern is given to word onset controls. The best control efforts amount to selecting only voiced initial words. A survey of naming data presented in the current literature was undertaken to examine the variability of naming times for different onset sound types. Naming data from Andrews (1989, 1992), Chumbley and Balota (1984), Waters and Seidenberg (1986).

A mean response time for each different voiced onset type from each study was determined and is shown in table 3–3. The general picture from the cross experiment survey is that approximants appear to activate the voice key much faster than stops or fricatives. Fricatives are slower than stops to trigger a response. In experiments 1, 2, 3, 4, 5, 7 and 8 of this thesis word onsets were carefully matched across conditions to ensure that naming effects were never artifacts of an uneven distribution of onset types.

3.1.6 Nonwords

Experiments 1a, 3a, 4a, and 8 involved the word/nonword discrimination task of lexical decision, while experiment 5 included a nonword naming condition. These experiments therefore required nonwords as well as real word stimulus sets. It is important to select nonword items carefully because in lexical decision subjects should reject nonwords because they do not have lexical entries, and not for some other reason induced by the way the nonwords are constructed. For example if all nonwords were completely odd looking orthographically illegal strings (e.g. *xbfk*) then subjects could make the decision on how weird the letters look and may never have to access the lexicon at all. In experiment 4 the wordlikeness of nonwords is manipulated, as was nonword neighbourhood size. In all other experiments nonword sets were matched to the word lists as closely as possible. Though matching word frequency between nonwords and real words is not appropriate, matching string length, orthographic neighbourhood size and onset letter was possible for all sets. Nonword strings were also chosen to be pronounceable so as to discourage subjects discriminating on the basis of pronunciation.

3.2 Tasks

The experimental tasks used in this thesis are lexical decision, naming and sentence reading. Lexical decision is believed to involve lexical access plus a decision process, whereas naming shares the access process but has a pronunciation process rather than a decision. The two tasks are often used in association since any variable which affects both tasks will be influencing lexical access whereas a variable which influences just one task may be affecting the process which is particular to that task. The sentence reading task involves neither a decision process, nor an overt pronunciation, it involves lexical access, and it will also involve higher level processes such as those concerning syntactic and semantic information. Any variable which influences all three tasks is therefore likely to be affecting lexical access. All three tasks are traditional word recognition tasks and are widely used and documented. For these reasons we will not discuss these tasks in general but will give details only of the specific task versions used in this thesis.

3.2.1 Lexical decision

A lexical decision task was used for experiments 1a, 2a, 4a, 7 and 8a. For experiment 4 *TSCOP* software run on a BBC was used, all other lexical decision experiments were executed by *PsyScope* software run on a Macintosh.

Instructions on the screen informed subjects that they would be presented with isolated four letter strings and that they should decide whether each string was a word or not a real word. They were instructed to indicate their decision by pressing the "WORD" or "NONWORD" button in front of them. Subjects rested their two fore-fingers on the response buttons, using their preferred hand for word responses. They were instructed to respond as quickly and as accurately as possible.

Stimuli were presented centre screen in black lower case letters against a white background. Each stimulus appeared until the subject had made a response, there was then a 1500 millisecond delay from the offset of one stimulus until the onset of the next. The screen was blank white during the inter-stimulus delay.

Ten practice trials were completed to familiarise the subject with the procedure before the experimental trials began. Subjects were given the opportunity to rest every forty trials or so. Response time and accuracy data (which button was pressed) were recorded by the computer.

3.2.2 Naming

A naming task was used for experiments 1b, 2b, 3, 4b, and 8. As for the lexical decision task, *TSCOP* software run on a BBC was used for experiment 4 and *PsyScope* on a Macintosh was used for all other naming studies. The stimuli were presented in exactly the same manner and format as for lexical decision. Instead of a lexical decision response subjects were instructed to read the word aloud as quickly and as carefully as they could. Responses were recorded from a throat microphone which the subjects wore above their larynx. Movement of the larynx at the very start of phonation triggered a pulse which recorded the time elapsed from the moment when the word appeared on the screen.

3.2.3 Sentence reading

Experiment 5 (see chapter 5) was the only one of the eight experiments in this thesis that did not involve isolated word recognition. Instead words were placed in neutral context sentences. Subjects were requested to read, in their own time, the sentences which appeared on the screen in front of them. Measurements were taken by a Dual-Purkinje eye tracking device for the times spent fixating on three regions of the sentence: the prior context of the target word, the target word, and the subsequent context of the target word.

When a subject reads text, eye movement may not be made purely left to right. The reader typically makes left to right saccades across the text. Second pass, regressive fixations may be made at any time, which involve re-inspection of the text in a right to left direction. The first pass is thought to be sensitive to word length, word frequency and syntactic information. The second pass is concerned with higher processing such as the checking of syntactic dependencies and the determining of antecedents of anaphoric expressions. It is therefore the first pass fixation times we examine as a response measure in Experiment 5.

To ensure that subjects were actually reading the sentence for comprehension a random selection of the sentences were followed by a question concerning the content of the sentence.

Chapter 4

Competition effects

4.1 The effect of a single strong competitor

As already stated in earlier Chapters the most commonly used metaphor to describe the process by which a single word is correctly identified from amongst all other possible candidates, is one of competition. When a word is perceived it is thought to generate a candidate set of possible alternatives. These candidate alternatives compete until a single word wins, and is thus recognised. However a candidate set of competitors is defined, it will contain words which resemble the target in some way (e.g. cohort members, neighbours, friends and enemies). Within a candidate set not every competitor will be equally strong. The strength of competitors is likely to depend on similarity to the target and on frequency relative to the target. A high frequency candidate which is very similar to a target will, therefore, be a very strong competitor. For visually presented words, similarity to target may be defined as the proportion of shared visual features. For example *boot* and *boat* are more similar than *park* and *type*. Some pairs of words share all of their letters (e.g. *salt/slat, from/form*), we may therefore expect the lower frequency members of these high similarity pairs to experience interference from their higher frequency competitor. This expectation is not however the prediction made by some models of lexical processing.

In an interactive activation account the stimulus *slat* will activate the letter units *s*, *l*, *a* and *t*. The active letter units will pass on activation to words containing those letters (e.g. *slate, slap*). The word *slat* shares all of its letters with the word *salt*. If both of these words were activated by the evidence of their four constituent letters, then without further information the processing system would be unable to choose between them. To avoid this problem Rumelhart and McClelland (1986) assume that letter order information is encoded in letter units by means of a context sensitive coding system. Any single letter is encoded in the context of those letters preceding and following it, a coding unit referred to as a Wickelgram (as suggested by Wickelgren 1969, see Rumelhart and McClelland 1986). The letter *l* of *slat* is contained in the Wickelgram '*sla* whereas the *l* in *salt* is encoded in the Wickelgram *alt*. The presentation of *slat* will, therefore, not pass activation to *slat* because *salt* and *slat* share no Wickelgrams.

Nonetheless when nonwords are made from real words which have undergone letter transposition (e.g. *trian* from *train*), they are harder to reject in LDT than other nonwords (Chambers 1979; O'Conner and Forster 1981). Furthermore, Chambers has shown that real words like *minuet* which have real word anagrams *minute* take longer to recognise than frequency-matched controls. A Wickelgram based interactive activation model would have to include single letters as well as wickelgram encoders to accommodate this within word letter transposition data.

Few people would be willing to argue for Wickelgrams as a very realistic form of representation, they are simply a device. Though it is necessary to have some device of position specificity, even Rumelhart and McClelland (1986) have suggested that the Wickelgram system in the Interactive Activation model may need to be relaxed. Relaxing the position specificity device may involve allowing other sources of information to influence the weighting given to Wickelgram information. For example, consider the Chambers (1979) word pair *minuet/minute*. Chambers finds interference between these two words in subject recognition, but a strict Wickelgram system will never confuse them because *uet* and *ute* are different Wickelgrams. Such *minuet/minute* type word pairs do however share the Wickelgrams *min* and *inu*, the first of which is also the first syllable or in Taft's (1979) terms, the BOSS. The pair also share some semantic component by reference to smallness. The similarity between the members of the stimulus pairs used by Chambers is considerable, and in a relaxed Wickelgram based system this similarity may out-weigh the scope of the single distinguishing Wickelgram.

In the present experiments we conduct a more stringent test of Wickelgram based position-specific sensitivity by using short word pairs (e.g. *slat/salt, loin/lion*). These pairs have no internal bigrams in common, no semantic relationship to one another, and most importantly share no Wickelgrams at all. The stimuli of the present experiment are related to their letter sharing competitor at only the letter level of representation. Because the relationship between these high similarity pairs is similar to the relationship between Coltheart's orthographic neighbours we will refer to these word pairs as near non neighbours (NNN).

4.1.1 Experiment 1a: LDT

Design

In a repeated measures design subjects made LDT responses to 60 randomly presented stimuli. Stimuli were either words or nonwords. Half of the word items could be made into a higher frequency word by transposing the middle letters (e.g. the low frequency word *warp* becomes the higher frequency word *wrap*, when the middle two letters are transposed), and half could not (e.g. *wean*). Similarly half of the nonword items could be made into a real word by transposing the middle letters (e.g. *bolb* from *blob*) and half could not (e.g. *slan*).

Cells of design		Mean values	
Lexical status	Competitor	Word freq	Nhood
Word	NNN	13	7
	Control	14	7
Nonword	NNN	0	3
	Control	0	3

Table 4–1: Experiment 1: Mean values of stimulus variables.

Subjects

The subjects were sixteen undergraduate and post-graduate volunteers from the University of Edinburgh. All of the subjects were native English speakers. All subjects had normal, or corrected to normal vision.

Materials

Word stimuli The word stimuli were thirty low frequency four letter words. The mean word frequency of an item was 13.4. Half of the words could be made into a higher frequency word by transposing the middle two letters (e.g. *slat* has a word frequency of 5, while *salt* has a frequency of 25). We describe these words as having near-non-neighbours (NNNs).

The other half of the words could not be made into other words by transposing their middle letters. Words in the transposable and non-transposable groups were matched for onset letter, word frequency and neighbourhood size. For example, *bran* which can be transposed to *barn* has a word frequency of 1

	Condition				
	NNN		Control		
Response	RT	PE	RT	PE	Mean RT
Word	626	9.5	608	4.1	617
Nonword	711	6.2	686	3.3	699

Table 4–2: Experiment 1a: Mean LDT times in milliseconds (RT) and percent errors (PE) for subjects.

and a neighbourhood size of 6. Its control is *boar* which cannot be transposed to produce another word but has a frequency of 1 and a neighbourhood size of 6.

Nonwords The nonwords were thirty pronounceable four letter strings, fifteen of which could be made into a word by transposing the middle letters (e.g. *bolb* from the word *blob*). The remaining fifteen did not make real words when their middle letters were transposed (e.g. *slan*). Onset letter and neighbourhood size were matched across the transposable and non-transposable nonword conditions.

Mean word frequency and neighbourhood size values for all word and nonword conditions of this experiment are shown in table 4–1. A list of all NNN words, nonwords and controls is given in Appendix A.

Procedure

The experiment was run using PsyScope software on the Macintosh. Ten practice and sixty experimental trials were run using the LDT methodology described in Chapter 3.

4.1.2 Results

Subject and item response means were analysed using t-test analyses. Raw data was trimmed separately for subject and item analyses. For each subject upper and lower response time cut-off values were set by calculating the subjects means response time plus or minus two standard deviations. A subjects responses which were more, or less than, two standard deviations from the mean were replaced by the cut-off values. Similarly for the by-items analysis cut-off values of an item mean response time, plus or minus two standard deviations were calculated for each item. Item responses greater, or less than two standard deviations from the mean were replaced by the cut-off values. The means for subjects are shown in table 4–2. Word and nonword responses were analysed in separate sets of t-tests. For the by-subjects analysis the effect of NNN (present/not present) was analysed within subjects. For the by-items analysis NNN effect was analysed between items.

Analysis of word response times. The 18ms difference between NNN and control words was significant by subjects ($t_1 = 2.14, df = 15, p < 0.05$) and by item analysis ($t_2 = 2.16, df = 13, p = 0.05$). Response times to NNN words were longer than those for matched controls.

Analysis of nonword response times. The 25ms difference between NNN and control nonwords is larger than the difference between NNN words and controls. The difference was significant by both subject and item analysis ($t_1 = 2.23, df = 15, p = 0.05$; $t_2 = 2.30, df = 13, p < 0.05$). Nonwords with a NNN took longer to reject than control nonwords.

	Condition				
	NNN		Control		
	RT	PE	RT	PE	Mean RT
Word response	549	1.8	558	1.6	554

Table 4–3: Experiment 1b: Mean naming times in milliseconds (RT) and percent errors (PE) for subjects.

Analysis of error data. For words significantly more errors were made to NNN items than to their controls, ($t = 3.72, df = 15, p < 0.005$). Similarly for nonwords significantly more errors were produced in responses to NNN items than to controls, ($t = 1.86, df = 15, p < 0.05$).

4.1.3 Experiment 1b: Immediate naming

Design and materials

The materials and design were the same as for the LDT experiment, 1a, but only the real words were used.

Subjects

The subjects were 11 new individuals from the same subject population as used for the previous experiment.

Procedure

Five practise and thirty experimental trials were run using the general immediate naming procedure described in Chapter 3.

4.1.4 Results

Responses were trimmed and analysed in the same way as the word items in the LDT experiment. NNN words were named 9ms faster than their controls, but this difference was not significant by subject or item analysis, ($t_1 = 1.39, df = 10, p = 0.20$; $t_2 = 1.89, df = 13, p = 0.10$).

4.1.5 Summary of Experiment 1

The investigation of near-non-neighbour competitors showed that the processing of both word and nonword targets in LDT is delayed by the existence of a single visually similar higher frequency competitor. LDT responses to NNN items took longer and were more erroneous than responses to matched control items. The demonstration of a NNN effect shows that competition between words may operate in a non-position-specific process. The influence of NNN competitors, although robust for LDT responses, had no influence on naming latencies.

4.2 The effect of competitor set size

The larger a potential candidate set is, the more words a target will have to compete with to be recognised. Early research into word envelope effects showed that the more infrequent a word envelope is the more likely subjects are to correctly identify or guess the word. Very common word envelopes will have many possible envelope filling candidates and without further physical information it may be impossible for a subject to select the correct candidate. Although in normal reading further physical information is available, envelopes were believed to represent a very early perceptual stage of word processing.

If the aim of word recognition is to narrow down the competitor set to a single target as quickly as possible, consideration of word shape, at least for short words would appear not to be very useful. For example two word envelopes (those of *bone* and of *cane* account for 30% of all the four letter words.

In Chapter two it was suggested that consideration of word beginnings and ends may be a more likely early focus of attention than envelopes. We refer to these beginnings and ends of words as word frames. A frame such as *n—e* has a very large candidate set of completion possibilities, whereas *b—p* leaves very few completion options. Strings with very common frames should therefore experience more competition and so take longer to respond to than those with rare frames. In the following experiment we investigate the effects of word frame frequency on response times in a LDT and a naming task.

4.2.1 Experiment 2a: LDT

Design

In a repeated measures design sixteen subjects made LDT responses to 60 stimuli which were presented in a random order. The stimuli were either words or nonwords and had either very common or very rare outer letters.

Subjects

The subjects for this experiment were sixteen undergraduate and postgraduate volunteers, from the University of Edinburgh. All the subjects were native English speakers, and had normal or corrected to normal vision.

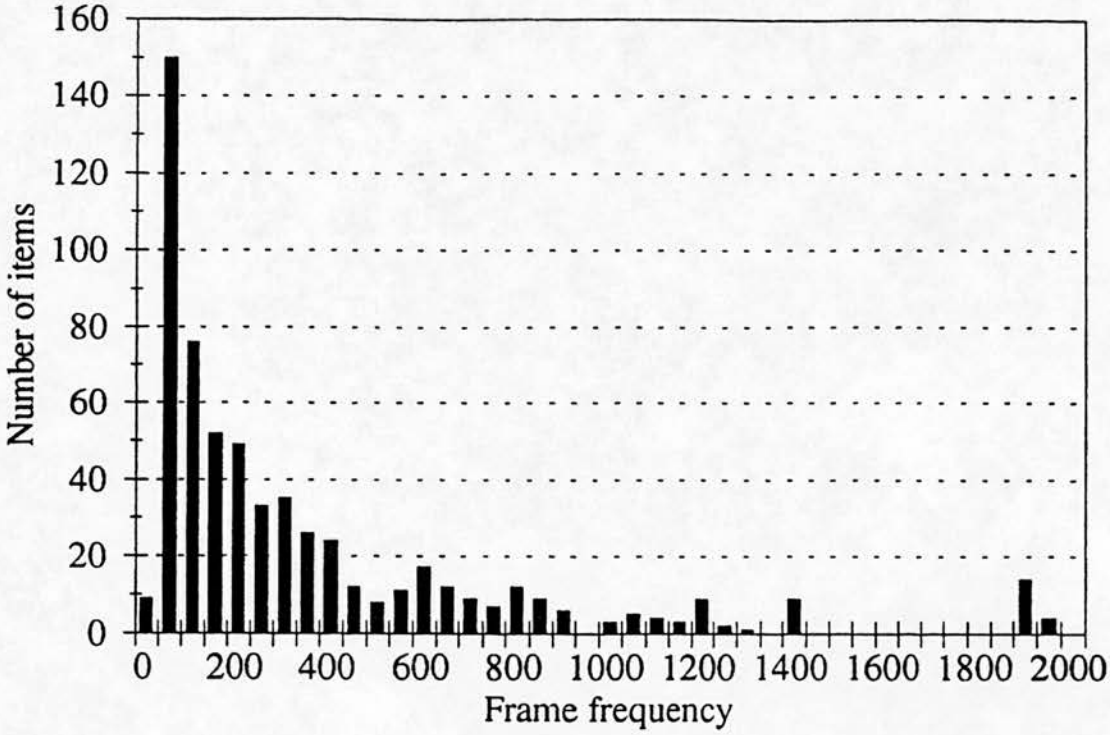

Figure 4–1: Frequency plot to show the distribution of word frames across a sample of 600 monosyllabic words.

Materials

Word stimuli The word stimuli were thirty low frequency four letter words. The mean word frequency was just 2.8, with a range from 0 to 10. Half of the words had very common outer letters: there were many occurrences of these initial and final letter pairs as the extremities of four letter words in the MRC Psycholinguistic Database, and half of the words had very rare outer letter pairs. We refer to the outer letter pairs as frames (see figure 4-1 for the distribution of word frames across a sample of 600 monosyllabic words).[1]

[1]A frame frequency value of zero means that the cumulative word frequency of all words sharing a particular frame is zero.

	Mean values			
	Targets		Frame sharers	
Frame	Word freq'	N'hood	No.	Word freq'
Common	2	6	12	421
Rare	3	5	1	0

Table 4–4: Experiment 2a: Mean values for stimulus variables.

The mean number of outer letter sharers for a common word frame was 12, the mean word frequency of outer letter sharers for an item in this group was 421. The rare frame items were either the only lexical entry of four letters having those particular extremities, or they shared outer letters with only one or two zero frequency words.

Onset letter, word frequency and orthographic neighbourhood size were matched across the two outer letter frequency conditions (see table 4–4 for mean values). For example the item *bunk* has common outer letters (e.g. *book, back, bank*), has a word frequency of 6 and a neighbourhood size of 10. This item was therefore matched with *bump* which has rare outer letters, a frequency of 5 and a neighbourhood size of 8.

Nonwords The nonwords for this experiment were thirty pronounceable four letter strings. Half the nonwords had outer letters commonly occurring for real words (e.g. *nule: name, node, nine, nape, note...*), and half had outer letters not found for any four-letter entry in the MRC database. The two nonword outer letter conditions were matched across condition for onset letter and neighbourhood size. The word and non-word stimuli for this experiment are shown in Appendix B.

	Frame				
	Common		Rare		
Response	RT	PE	RT	PE	Mean RT
Word response	614	4.5	588	5.0	601
Nonword response	705	5.4	672	2.6	689

Table 4–5: Experiment 2a: Mean LDT times in milliseconds (RT) and percent errors (PE) for subjects.

Procedure

The experiment was run using PsyScope software on a Macintosh computer. Ten practise and sixty experimental trials were run using the LDT procedure described in Chapter 3.

4.2.2 Results

Mean response times of subjects and of items were analysed with t-test analyses. Response times of a subject that were less than, or greater than two standard deviations from the mean of that subject were replaced by those cutoff values. Similarly response times to an item that were less than, or greater than two standard deviations from the mean of that item were replaced by those cutoff values. Separate analyses were performed for word and non-word responses, with both subject and item means. For the by-subjects analysis the effect of word frame was examined within subjects. For the by-items analysis word frame was analysed between items.

	Frame				
	Common		Rare		
	RT	PE	RT	PE	Mean RT
Word response	549	2.7	539	1.3	544

Table 4–6: Experiment 2b: Mean naming times in milliseconds (RT) and percent errors (PE) for subjects.

Analysis of word response times. Words with common outer letter pairs were judged on average 25ms more slowly than matched words with rare outer letters. This difference was significant by subject ($t_1 = 2.66, df = 15$, $p < 0.02$, and by item $t_2 = 2.26, df = 13$, $p = 0.05$) analyses.

Analysis of nonword response times. Nonwords with rare outer letters were rejected 33ms faster than nonwords with lexically common outer letters. For nonwords the rare outer letter advantage was significant by subject ($t_1 = 3.88$, $df = 15$, $p < 0.01$) and item ($t_2 = 2.56$, $df = 13$, $p < 0.05$) analyses.

Analysis of error data. There were no significant differences in any error rates.

4.2.3 Experiment 2b: Immediate naming

Design and Materials

The experimental design and materials were the same as those used for Experiment 2a, except that for the naming task only the word stimuli were used.

Subjects

The subjects were eleven new individuals from the same population as used in Experiment 2a.

Procedure

The procedure followed exactly that described for naming in Chapter 3.

4.2.4 Results

The data were trimmed and analysed in the same way as the word data of Experiment 2a. Subject mean response times are shown in table 4–6. The 10ms difference between common and rare frame words was not significant in either subject or item analysis, ($t_1 = 1.42, df = 10$, $p = 0.20$; $t_2 = 1.10, df = 13$, $p = 0.20$).

4.2.5 Summary of Experiment 2

Nonwords appear to be much easier to reject in LDT if they have rare outer letter frames which do not occur for real lexical items. If word frames are encoded at an early stage of recognition, then a frame that has no lexical representation can quickly be rejected as belonging to a nonword. For real words having a great many letter frame competitors slows LDT responses. This effect of competition is not found for the naming task.

4.3 Competition at word onset

Experiments 1 and 2 show that large candidate sets result in strong competition, and that a single high frequency candidate, visually similar to the target, is a very strong competitor. Both of these competition effects are manifest only in the LDT task and not in word naming. The definition of a NNN competitor or a word frame is made in terms of orthography. LDT responses may involve purely orthographic processing, mapping the letters of the visual stimulus onto letter representations in the lexicon. Unlike LDT, naming must involve phonological processing at some stage, so as to produce a pronunciation. It may be, therefore, that the candidates invoked in a naming task are related to the phonology of the target word. Since the naming task involves naming of visually presented words we may expect the mapping of orthography to phonology to be especially important in determining the candidate set.

In the task of word naming, latencies are recorded from word onset. The phonological code of the initial diphone, if not the whole word, must be retrieved or constructed by the time of response. There are at least two alternative ways in which people may perform the naming task. Firstly subjects may form a phonological code after some global consideration of the whole word. Secondly subjects may produce the phonological onset code before consideration of the word, so as to trigger a very fast response (Taft 1991).

Whichever of the two alternatives is correct, the speed of phonological assignment may depend on the reliability of the onset's spelling-sound relationship. A candidate set may be invoked by onset letters, and competition may then occur between the pronunciation alternatives within the active candidate set. For example, the visual stimulus *dog* may activate the digram sharing candidates *dog,* *dot, doe, don*, but it may be the diphones */d@U/* and */d0/* which compete[2]. We may expect therefore that words with very common, or consistent, onset digram to diphone correspondences at onset should be named more quickly than those with rarer, or irregular correspondences. In the following experiment we investigate the effect of onset digram to diphone correspondence consistency on the speed of immediate naming. This variable concerns the mapping of orthography to phonology. We also consider the effect of orthographic neighbourhood size (Coltheart et al. 1977), on naming responses. Neighbourhood size is a purely orthographic variable and we therefore expect that like NNNs and word frames it will not influence naming.

4.3.1 Experiment 3

Design

Subjects made naming responses to one hundred and twenty eight target and filler items presented in a random order. The sixty four target words filled a 2 $\times$ 2 $\times$ 2 manipulation of onset digram to diphone correspondence consistency (consistent/non-consistent), word frequency (high/low), and neighbourhood size (large/small).

Subjects

The subjects were twenty new individuals from the population described in Experiment 1.

---

[2]These phonetic transcriptions are taken from the MRC Psycholinguistic database, see Coltheart, (1981).

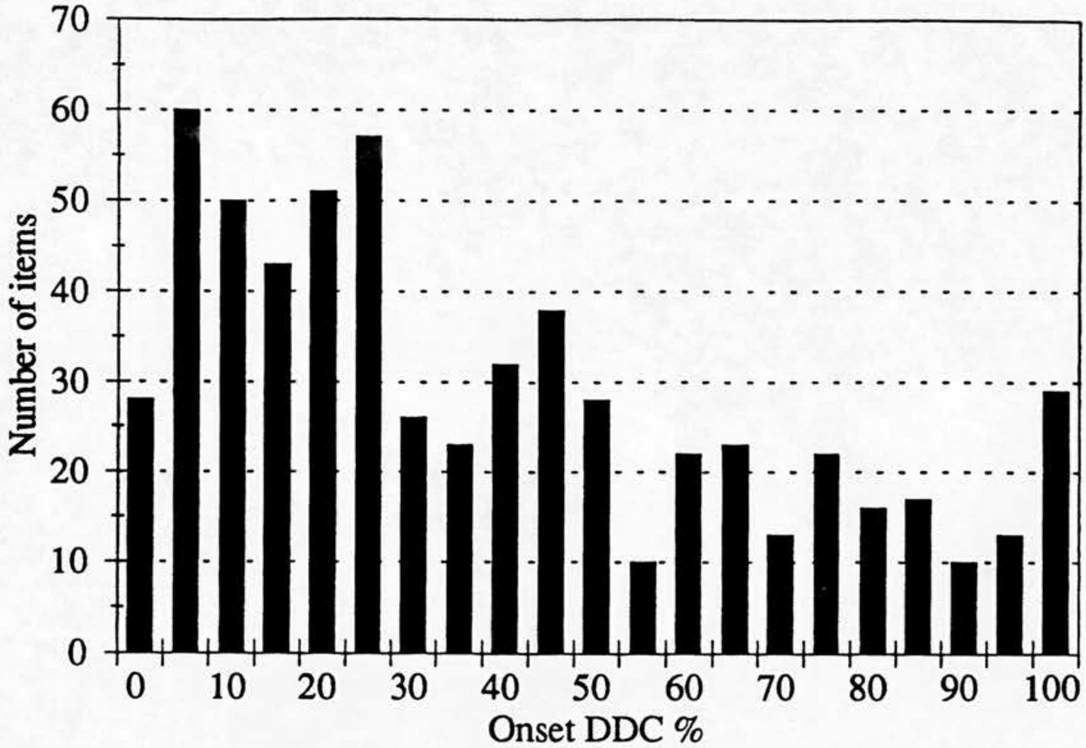

Figure 4–2: Frequency plot to show the distribution of onset DDCs across a sample of 600 monosyllabic words.

Materials

The targets were sixty-four, four-letter words. Half of the words had consistent onset digram to diphone correspondences (DDCs), and half had non-consistent onset DDCs. To determine onset DDC consistency we considered a sample of all the monosyllabic words in the MRC Psycholinguistic database. For each onset digram, we found the total word frequency of all entries sharing that onset digram. We then took each onset diphone occurring within an onset digram set, and determined the percentage of total word frequency accounted for by each onset diphone. The onset DDC therefore represents a frequency based measure of the percentage of words sharing a targets onset digram which also share its onset diphone. For example 40% of occurrences of lexical items which begin with

Cells of design			Mean values		
Onset diphone	Word freq	Nhood	Word freq	Nhood	Onset
Consistent	High	Large	65	13	54
		Small	68	6	54
	Low	Large	3	15	54
		Small	2	4	54
Non-consistent	High	Large	77	13	19
		Small	31	6	19
	Low	Large	3	14	19
		Small	2	4	19

Table 4–7: Experiment 3: Mean values for stimulus variables.

the letters *do* start with the diphone */d@U/*. Figure 4–2 shows the distribution of onset DDC consistency for 600 monosyllabic words.[3]

For words with consistent onset DDCs, about 60% of occurrences of that onset digram gave rise to that diphone. The non-consistent onset DDCs were shared by less than 20% of words starting with the target letters. Within each of the onset DDC conditions, words were split equally into high and low frequency groups. High frequency words had a mean word frequency of 60; low frequency words had a mean of 3. Each of the four onset $\times$ word frequency conditions were further divided into large or small neighbourhood size conditions. Large

		Onset diphone				
		Consistent		Non-consistent		
Word freq	Nhood	RT	PE	RT	PE	Mean RT
High	Large	458	4.5	469	6.0	464
	Small	455	7.0	471	4.0	463
High freq means		457	5.7	470	5.0	464
Low	Large	474	6.5	480	2.0	477
	Small	469	6.0	482	6.0	476
Low freq means		472	6.2	481	4.0	477

Table 4–8: Experiment 3: Mean naming times in milliseconds (RT) and percent errors (PE) for subjects.

neighbourhoods had a mean of 14 neighbours, while small neighbourhoods had a mean size of 5 (see table 4–7 for mean values).

Across the four word frequency × neighbourhood size cells of each of the two onset groups, onset diphone was matched. A complete list of the target words is provided in Appendix C.

Sixty-four filler words were also included in the stimulus set to increase the number of different onset diphones being used.

Procedure

The experiment was run on the Macintosh using PsyScope software. Subjects were given ten practice trials before one hundred and twenty eight experimental trials. There were two evenly spaced opportunities to rest during the experi-

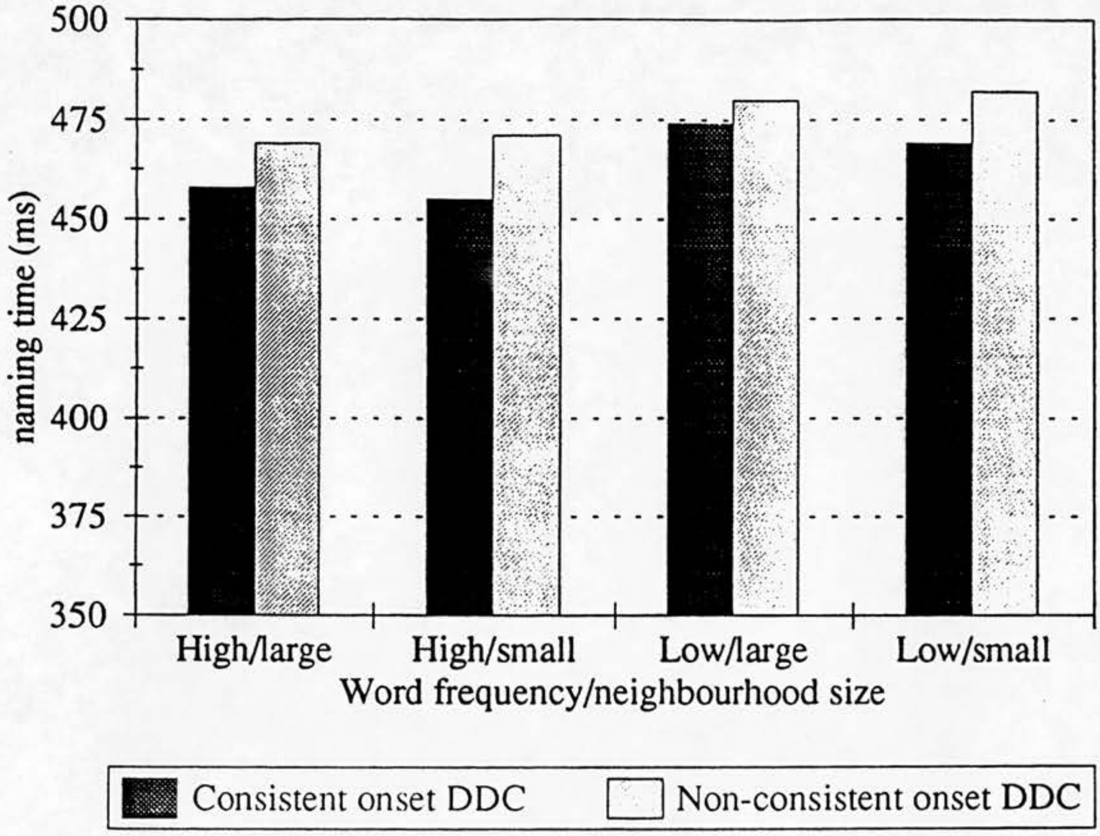

Figure 4–3: Mean naming latencies (ms) of the word frequency × neighbourhood size conditions for each onset DDC condition in Experiment 3.

mental trials. The procedure followed the immediate naming methodology, as described in Chapter 3.

4.3.2 Results

Mean naming latencies are shown in table 4–8. The data were trimmed in the same way as the data of Experiment 1. Subject and item means were analysed with a 2(onset digram to diphone correspondence: consistent/non-consistent) × 2(word frequency: high/low) × 2(neighbourhood size: large/small) ANOVA. For the by-subjects analysis all three factors (onset DDC, word frequency and neighbourhood size) were within subject factors. For the by-item analysis all three factors were analysed between items.

The 13ms advantage for high frequency words over low frequency words was significant by *min*F' analysis ($F_1(1,19) = 20.43$, $p < 0.0002$; $F_2(1,56) = 4.91$, $p < 0.03$; $minF'(1,74) = 3.96, p < 0.05$). Words with consistent onset DDCs were named 11.88ms faster than those with non-consistent onset DDCs. This effect of onset DDC consistency was significant by both subject and item analysis, but only marginal in the *min*F' analysis ($F_1(4,88) = 9.43$, $p < 0.006$; $F_2(1,50) = 3.97$, $p < 0.05$; $minF'(1,74) = 2.79, p = 0.07$). Orthographic neighbourhood size had no effect on naming times ($F_1 < 1; F_2 < 1$) (see figure 4–3). None of the interactions were significant (all $F < 1$).

4.3.3 Summary of Experiment 3

Naming latencies were facilitated by word frequency: high frequency words were named faster than low frequency words. Words with consistent digram to diphone correspondences (DDCs) at onset were named faster than those with non-consistent onset DDCs. Orthographic neighbourhood size had no effect on naming times. Neighbourhood size is a purely orthographic measure, whereas onset DDC is related to the word's pronunciation. It would appear therefore that variables relevant to phonological processing are more likely to be influential in naming than orthographic variables.

[3] A value of zero means that less than 1% of occurrences of a certain onset digram share the same onset diphone, a very non-consistent onset DDC.

4.4 Discussion

Experiments 1 and 2 both show some form of competition effect for LDT and not for naming. Experiment 3 shows an effect for naming based on the manipulation of a phonologically related variable, but no effect for orthographic neighbourhood size. NNNs, word frames and neighbourhoods all concern orthography and none shows an effect on naming. It would seem, therefore, that to influence naming, a variable must capture some aspect of the orthography to phonology mapping which is necessary for the print to speech response. Taken together the experimental results of this Chapter suggest that the nature of a word recognition task may determine the candidate set which is activated, or the candidates which compete.

4.4.1 Competition between candidates

Evidence from the experiments presented in this Chapter suggest two related situations in which the presence of competitors can delay a target's recognition;

1. When there is a single very strong non-target competitor in the candidate set.

2. When the candidate set is very large.

The strongest competitors in a candidate set will be those items that; a) are most similar to the target, and b) have a high frequency of occurrence relative to the target. NNNs are therefore very strong competitors on the grounds of both similarity and frequency. Experiment 1 showed that low frequency members of NNN pairs are responded to more slowly in LDT than carefully matched controls.

The effects on target recognition of a higher frequency NNN, appear similar to the reported effects of a single higher frequency orthographic neighbour. In situations where neighbourhood size has been shown to delay the recognition of low frequency words, a single higher frequency neighbour is just as inhibitory as several higher frequency neighbours, regardless of the number of lower frequency neighbours (Grainger, O'Regan, Jacobs and Segui 1989; Grainger 1990; Grainger and Segui 1990).

Bard and Shillcock (1992) have shown that the distribution of frequencies within lexical competitor sets is such that most contain a single high frequency outlier. This outlier they have shown to have the maximum advantage in frequency based activation over any other word in the competitor set. The frequency advantage available to any other competitor will be small in comparison, and so there is typically only ever a single really strong competitor in a candidate set. Furthermore Bard and Shillcock have shown that bigger competitor sets contain more frequent words: however the candidate set is formulated, size is positively correlated with the frequency of the most frequent competitor.

The consideration of word frames in Experiment 2 allowed the definition of competitors to cover quite large sets of lexical items. The finding from Experiment 2 was that words with very common outer letters are responded to more slowly than those with rare outer letter frames. Given the correlation between competitor set size and frequency of the most frequent competitor, competitor sets defined as common word frames were more likely to include a very strong non-target competitor than those defined as having rare frames.

4.4.2 Conspiracy between candidates

Experiment 3 showed that naming responses are sensitive to the consistency of digram to diphone correspondence (DDC) at word onset. Words with very consistent onset DDCs were named faster than those with non-consistent DDCs.

The more tokens which begin with a certain onset DDC, the more quickly that onset DDC is produced. Unlike the NNN and word frame effects, the size of the onset DDC candidate set does not adversely affect subjects responses.

One reason for the beneficial influence of onset DDC consistency may be because our definition of a consistent onset DDC selected items having the most frequent onset DDC of the possible alternatives. These frequent DDCs are more likely to have only lower frequency DDC competitors and so may experience less competition than non-consistent DDCs which are likely to have higher frequency DDC competitors.

The onset DDC effect is similar to the word body findings of Jared, McRae and Siedenberg (1991). Friends are words sharing a target's body pronunciation (e.g. *sown: own, grown*). Enemies are words which share a target's body letters but do not share pronunciation (*sown: crown, down*). Words which have a greater cumulative frequency of friends than enemies are more rapidly named than words with more frequent enemies than friends. Words with consistent onset DDC's can be thought of as those with a higher percentage of onset friends than enemies. Similarly words with rare non-consistent DDCs are those with more enemy than friend onsets.

Both NNN and word frame candidates compete, but it is unclear that competition is necessary for the onset DDC effect. The definitions of NNN and word frame consider whole word level representations. In contrast the onset DDC is a word part. While word level representation may have to compete for a single winner to be recognised, it may be that representations of the constituent parts of a word can conspire. NNN and word frame effects are found only in LDT and not for naming, whereas onset DDC effects do emerge in naming. Lexical decision is considered to be more reliant on lexical processing than naming, which it is thought can be performed lexically or sublexically. The results of Experiments 1, 2 and 3 may therefore suggest two hypotheses; firstly, that word representations compete, while sublexical word parts conspire, and secondly that the task subjects are asked to perform determines the type of processing they employ and thus the candidate sets they consider.

4.4.3 Position Specificity

The finding in Experiment 1 that low frequency members of NNN pairs are responded to more slowly in LDT than matched controls suggests not only that more frequent than target, letter sharing competitors are very strong competitors, but also that similarity between competitors need not be position specific. This NNN result of non-position specific competition elegantly demonstrates, for example, the inadequacy of the Wickelgram unit of representation. Competition was found between words which had no Wickelgram overlap at all.

Although Wickelgrams are certainly not the right device for position specificity in word recognition, some form of position sensitivity is necessary. Experiment 2 showed that the word recognition system has some sensitivity to the outer letter frames of words. Frames are very fixed position competitors since only if the letters are at the beginnings and ends of words can they constitute a frame. The outer letters being more surrounded by white space on the page than intervening letters may be easier to specify positionally than those intervening letters. Furthermore the frames of short words are often consonants while the middles consist of vowels. This makes the word edges more informative than the more redundant vowel dominated centres. It may be therefore that non-position-specific competition is particularly relevant to the middles of words, and that position specificity in visual word recognition is especially relevant to the edges of words.

Chapter 5

Orthographic neighbourhood size

5.1 Orthographic neighbourhood size effects in LDT and naming

In the previous Chapter we demonstrated that the task subjects are asked to perform (LDT or naming) may to some extent determine which candidates compete for recognition. For the naming task competition effects were only found for the onset digram to diphone consistency variable and not for any of the purely orthographic measures (NNN, word frame, neighbourhood size). We also showed that competition from a single strong candidate (NNN) or from a large candidate set (word frame) delays LDT responses.

The results of Andrews (1989,1992) contrast dramatically with the findings of the previous Chapter. Andrews reports that large competitor sets can be facilitatory, and that orthographically defined candidate sets in the form of neighbourhoods can affect naming responses as well as LDT. The disparity between the findings of Andrews and the conclusions of Chapter 4 demands further investigation.

In the following experiment we repeat the experiments of Andrews. For this repetition we have made two enhancements to the original studies. Firstly we have used a more tightly controlled stimulus list. Secondly we have included all of the investigated variables (word frequency, neighbourhood size, onset, and

Cells of design			Mean values		
Lexical status	Word freq	Nhood size	Word freq	Nhood size	Bigram freq
Words	High	Large	118	20	5799
		Small	129	8	5563
	Low	Large	3	21	5681
		Small	3	8	5668
Wordlike nonwords		Large		20	5934
		Small		7	5677
Unwordlike nonwords		Small		6	5577

Table 5–1: Average word frequency, neighbourhood size, and bigram frequency for the word and nonword materials of Experiment 4

bigram frequency) in a single experiment, rather than having different variables examined in separate experiments.

5.1.1 Experiment 4a: LDT

Design

The experiment was a 2 × 2 manipulation of word frequency (high/low) and neighbourhood size (large/small). Bigram frequency and word onset were controlled across the four conditions. Twenty-four subjects gave LDT responses to 40 words and 40 nonwords. The presentation order was random. Twelve subjects saw wordlike nonwords and 12 saw unwordlike-nonwords. The wordlike nonwords were further divided into large and small neighbourhood size conditions.

Subjects

The subjects were 24 new individuals from the same population as described for the previous experiments.

Materials

Word Stimuli. The word stimuli were 40 four-letter English nouns. They were selected from the M.R.C. Psycholinguistic Data Base to conform to a 2 × 2 factorial design where the two factors were word frequency (high/low) and neighbourhood size (large/small).

High frequency words were those appearing in more than 35 samples according to the Kučera and Francis (1967) norms, and those classed as low frequency appeared in less than 9 samples. Words classed as having large neighbourhoods were those with at least 16 neighbours and those classed as small had less than 11. All four cells were matched for total bigram frequency. The two neighbourhood size conditions were matched for word frequency, and the two frequency conditions were matched for neighbourhood size. Descriptive statistics for the stimuli in each condition are summarised in table 5–1.

Many of the words in Andrews' stimulus set were neighbours of one another for example; *mine*, *mire* and *mite*, all from the same neighbourhood. It may have been possible for neighbours to act as primes and facilitate or inhibit one another's' recognition. To avoid this sort of interference among stimuli, no stimulus in the set constructed for the present experiment was a neighbour of any other stimulus. This constraint extended also to non-word stimuli.

In the Andrews words there exist homophones like *taut/taught*. These words may be more complex to process than non-homophonic words. For this reason

High/Large	High/Small	Low/Large	Low/Small
give	gift	gilt	glue

Table 5–2: Example of an unacceptable set.

only words with a single orthography to phonology correspondence were included in the materials.

Word initial diphone was controlled across the four cells of the design, no word in any letter onset group had a unique initial diphone. For example the sequence in Table 5–2 would not be acceptable as *glue* differs in initial diphone from its controls. This control also balanced the number of voiced and voiceless onsets between cells. All the word stimuli were nouns. That is to say that they had, as a minimum requirement, a noun entry in the dictionary, though some of the stimuli have other form class entries.

Non-Word Stimuli. As in the original Andrews (1989) experiments there were two different nonword conditions within the design, wordlike non-words, and unwordlike non-words. The more unwordlike nonwords are, the easier it will be to distinguish them form real words, and thus the easier the decision task will be. Manipulating the wordlikeness of nonwords should not however influence the access process, since whatever the nonwords are like, it is only the words which have lexical entries. Any effects which truly influence the lexical access process, and are not just artifacts of the decision process, should therefore be unaffected by the manipulation of nonword wordlikeness.

For the wordlike condition, forty wordlike non-words of four letters were constructed by changing one letter of a real word to produce a letter string that was still legal and pronounceable, but was not a real word or a homophone of a real word. The wordlike nonwords used here were in two sets, those with large neighbourhoods and those with small neighbourhoods. For the unwordlike nonwords this division was not possible. Unwordlike nonwords were those having

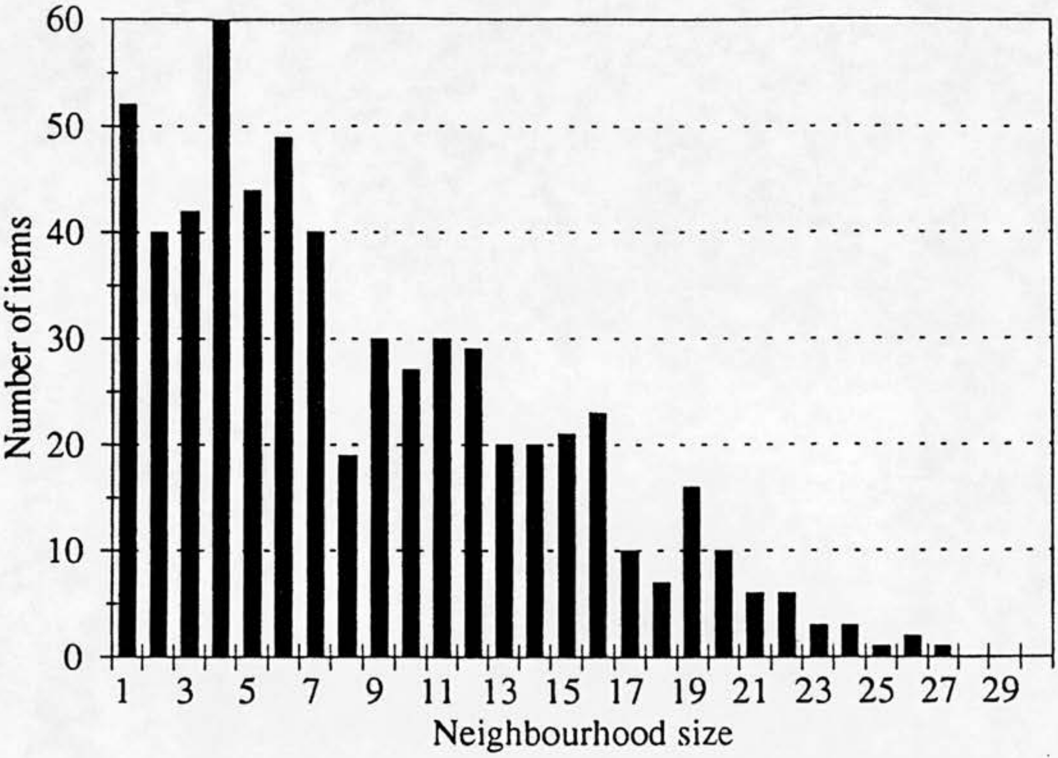

Figure 5–1: Frequency plot to show the distribution of neighbourhood size across a sample of 600 monosyllabic words.

illegal or very rare letter clusters. Once a four letter word has such a rare cluster within it, the number of neighbours that can be made by altering one letter of that word becomes quite small, especially when there is the added constraint that the bigram frequency mean of the group has to be kept equivalent to those of the other stimulus conditions.

Twenty of the wordlike non-words had large neighbourhoods with a mean size of 19.5, and twenty had small neighbourhoods with a mean size of 7.35 neighbours. The neighbourhood sizes for wordlike non-words were therefore equivalent to those of the word stimuli. (See figure 5–1 for the range of monosyllabic word neighbourhood sizes.)

Bigram frequencies for large and small neighbourhoods of wordlike non-words were also kept equivalent to those of the word stimuli (see figure 5–2 for the distribution of word bigram frequency within monosyllabic lexical entries). The

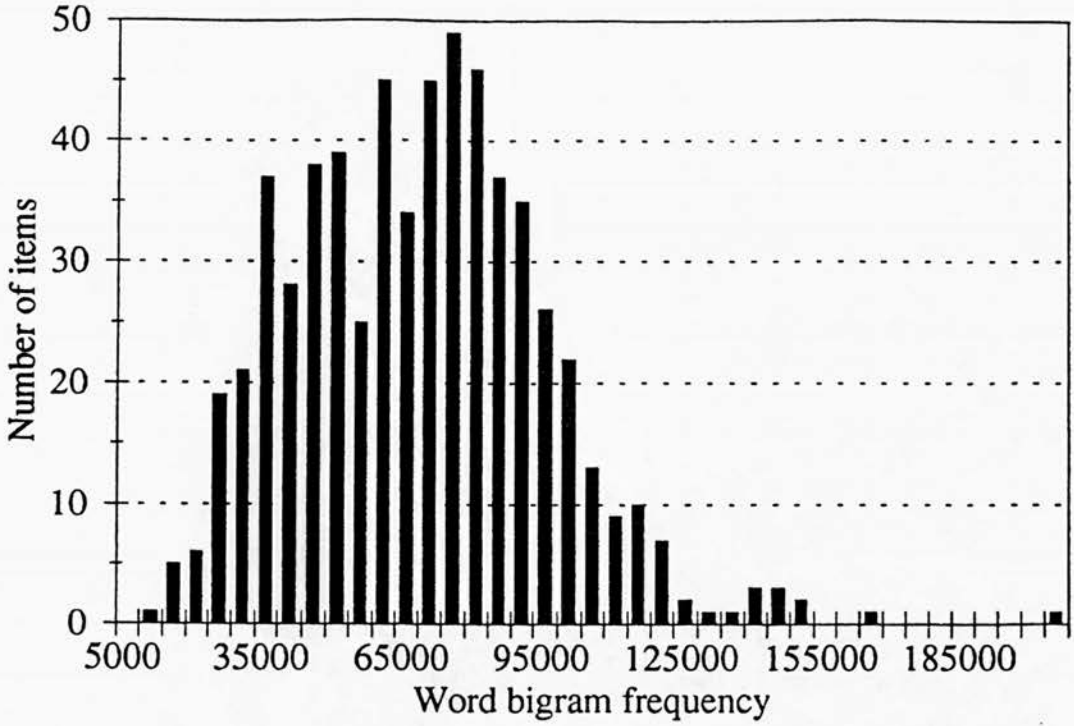

Figure 5–2: Frequency plot to show the distribution of bigram frequency across a sample of 600 monosyllabic words.

means being; 5934.11 for stimuli with large neighbourhoods, and 5677.4 for those with small neighbourhoods.

For the unwordlike non-word condition, forty four-letter nonwords were created that contained either an illegal cluster of letters or a legal but very rare cluster, rare at least in that position or size of word. This made the unwordlike non-words crucially less wordlike that those of the wordlike condition.

The mean neighbourhood size of these forty unwordlike non-words was 5.5. Bigram frequency was again maintained at a level comparable with the word stimuli, with a mean of 5577.405. [1] A full list of word and nonword stimuli used for this experiment is given in Appendix D.

---

[1]If bigram frequencies for non-words are equivalent to those of words, then the non-words may not be truly unwordlike. Rare and illegal clusters have very rare bigram

		Neighbourhood size				
		Large		Small		
Response	Freq	RT	PE	RT	PE	Mean RT
Words: wordlike foils	High	540	0.0	544	0.0	542
	Low	565	6.6	579	6.6	572
Words: unwordlike foils	High	504	0.0	498	0.3	501
	Low	518	3.3	557	4.1	538
Nonwords		644	2.5	611	1.2	628

Table 5–3: Experiment 4a: Mean lexical decision times in milliseconds (RT) and percent errors for subjects

Procedure

Subjects first completed 16 practice trials and then received 80 experimental trials preceded by two buffer items. Instructions and stimuli were presented on a V.D.U. screen controlled by a BBC microcomputer that also recorded response latencies and errors. The procedure followed that of the general LDT methodology, as described in Chapter 3.

5.1.2 Results

Data points more than two standard deviations from the mean were trimmed as described for Experiment 1. Mean lexical decision times for subjects and for items were analysed by a 2(wordlikeness of foil: wordlike/unwordlike) ×

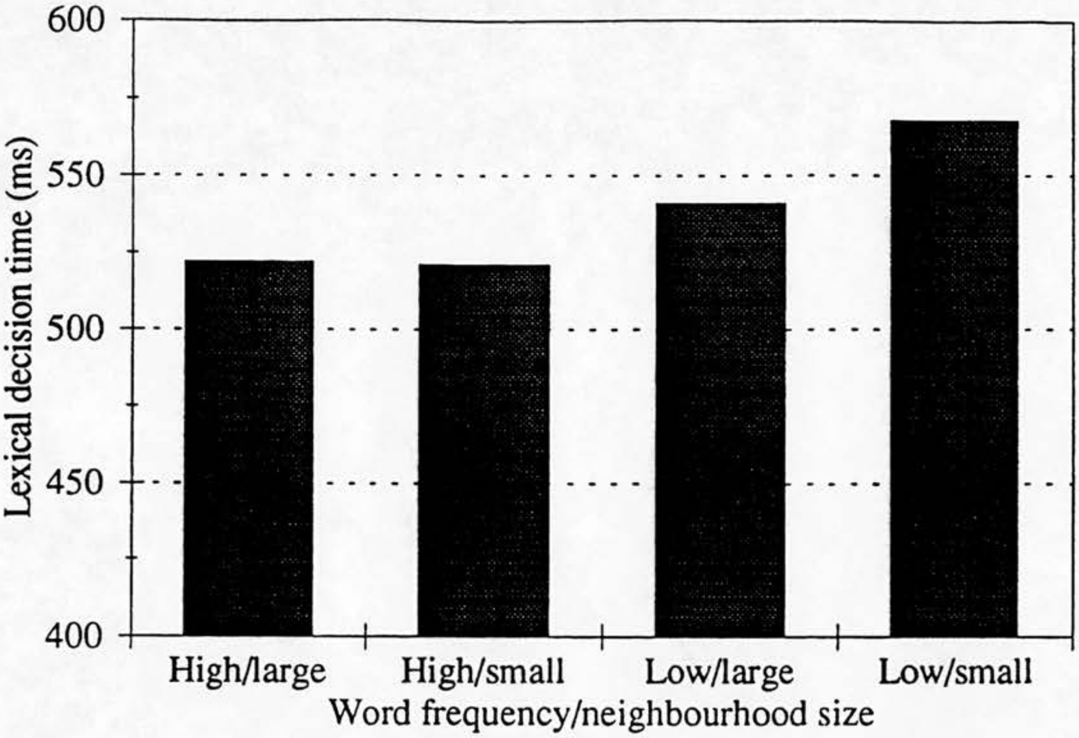

Figure 5–3: Mean lexical decision times (ms) for the word frequency × neighbourhood size conditions of Experiment 4a.

2(word frequency: high/low) × 2(neighbourhood size: large/small) ANOVA. For the by-subjects analysis wordlikeness was a between subjects factor whereas frequency and neighbourhood size were both within subject variables. For the by-items analysis word frequency and neighbourhood size were between item factors whereas wordlikeness was a within items variable. Significant interactions were further examined with Scheffé tests.

Analysis of word response time. Analysis of the LDT data revealed a main effect of word frequency that was significant by $minF'$ ($F_1(1,22) = 30.77$, $p < 0.00001$; $F_2(1,36) = 24.20$, $p < 0.0001$; $minF'(1,57) = 13.55$, $p < 0.01$): higher frequency words were responded to faster than lower frequency words. The effect of neighbourhood size was found to be significant by subjects but only marginal by items, ($F_1(1,22) = 5.92$, $p < 0.02$; $F_2(1,36) = 3.44$, $p < 0.07$):

words from larger neighbourhoods were responded to more quickly than those from smaller neighbourhoods.

The interaction between word frequency and neighbourhood (see figure 5–3) size was significant both by subject and item analyses, but only marginal by $minF'$ analysis ($F_1(1,22) = 4.59$, $p < 0.043$; $F_2(1,36) = 3.98$, $p < 0.053$; $minF'(1,56) = 2.13, p = 0.08$). Scheffé tests showed that low frequency words were responded to more quickly if they had large orthographic neighbourhoods ($F_1(1,22) = 8.74, p < 0.01$; $F_2(1,36) = 9.08, p < 0.01$), but that high frequency words were not affected by neighbourhood size ($F_1 < 1 : F_2 < 1$).

The effect of wordlikeness of nonword stimuli was significant by items and not by subjects,($F_1(1,22) = 2.12$, $p < 0.159$, $F_2(1.36) = 30.77$, $p < 0.00001$): word responses were made more quickly when nonword stimuli were more unwordlike. Wordlikeness did not interact significantly with any other variable. The effects of word frequency and neighbourhood size are not changed significantly by making the words more unwordlike, and are therefore likely to be influencing lexical access rather than decision processes.

Analysis of nonword response time. Wordlike nonwords showed a significant effect of neighbourhood size by $minF'$ analysis ($F_1(1,11) = 11.55, p < 0.01$, $F_2(1,19) = 9.89$, $p < 0.01$; $minF'(1,28) = 5.33, p < 0.05$), nonwords with large neighbourhoods took longer to respond to than nonwords with fewer neighbours. For the nonwords we have, therefore, replicated the findings of Coltheart, Davelaar, Jonasson and Besner (1977).

---

frequencies. To maintain total bigram frequency scores at a level comparable to word conditions, at least one very high frequency bigram must be included. This procedure satisfies a bigram control measure, but if trigrams were considered a total score for the unwordlike nonword condition would be lower than for word conditions.

Analysis of error data. In the analysis of errors, the only significant result was a main effect of word frequency, ($F_1(1,22) = 12.04$, $p < 0.002$, $F_2(1,36) = 11.86$, $p < 0.001$; $minF'(1,55) = 5.97, p < 0.05$), more errors were made in responses to low frequency words than to higher frequency words.

5.1.3 Experiment 4b: Immediate and delayed naming

Design and Materials

The experiment was a 2 × 2 manipulation of word frequency (high/low) and neighbourhood size (large/small). Twelve subjects were asked to name the words as soon as they possibly could and twelve were asked to wait until a set of brackets appeared. The 40 words from the previous LDT experiment were used.

Subjects

Twenty-four new subjects from the same subject population as used for the previous experiment were used for the present investigation.

Procedure

Using the same equipment set up as in the LDT experiment, 8 practice and 40 experimental trials were run. Half of the subjects performed immediate naming under the same procedure as described in Chapter 3. The other half of the subjects performed delayed naming. Presentation of stimuli for delayed naming was the same as for immediate naming and LDT. In delayed naming conditions subjects were told not to pronounce the word until a set of brackets appeared on the screen surrounding the word. To avoid reliable anticipation of the brackets, the delay between stimulus presentation and the appearance of the brackets varied randomly between intervals of 500, 800, and 1100 milliseconds. Responses were recorded in the same way as those in immediate naming.

		Neighbourhood size				
		Large		Small		
Naming Pt	Freq	RT	PE	RT	PE	Mean RT
Immediate	High	565	1.6	576	2.5	571
	Low	586	5.0	595	1.6	591
Delayed	High	422	5.8	438	2.5	430
	Low	423	2.5	416	4.1	420
mean		499	3.7	506	2.6	503

Table 5–4: Experiment 4b: Mean naming response times in milliseconds (RT) and percent errors (PE) for subjects.

5.1.4 Results

Naming latencies for both subjects and items were trimmed in the manner described for Experiment 1. Both subject and item mean latencies were analysed by a 2(naming point: immediate/delayed) × 2(word frequency: high/low) × 2(neighbourhood size: large/small) ANOVA and Scheffé tests. For the by-subjects analysis word frequency and neighbourhood size were analysed within subjects whereas naming point was a between subjects factor. For the by-items analysis naming point was analysed within items whereas word frequency and neighbourhood size were between item factors.

The effect of naming point (immediate or delayed naming) was significant by $minF'$ analysis, ($F_1(1, 22) = 30.40$, $p < 0.0001$; $F_2(1, 36) = 684.29$, $p < 0.0001$;

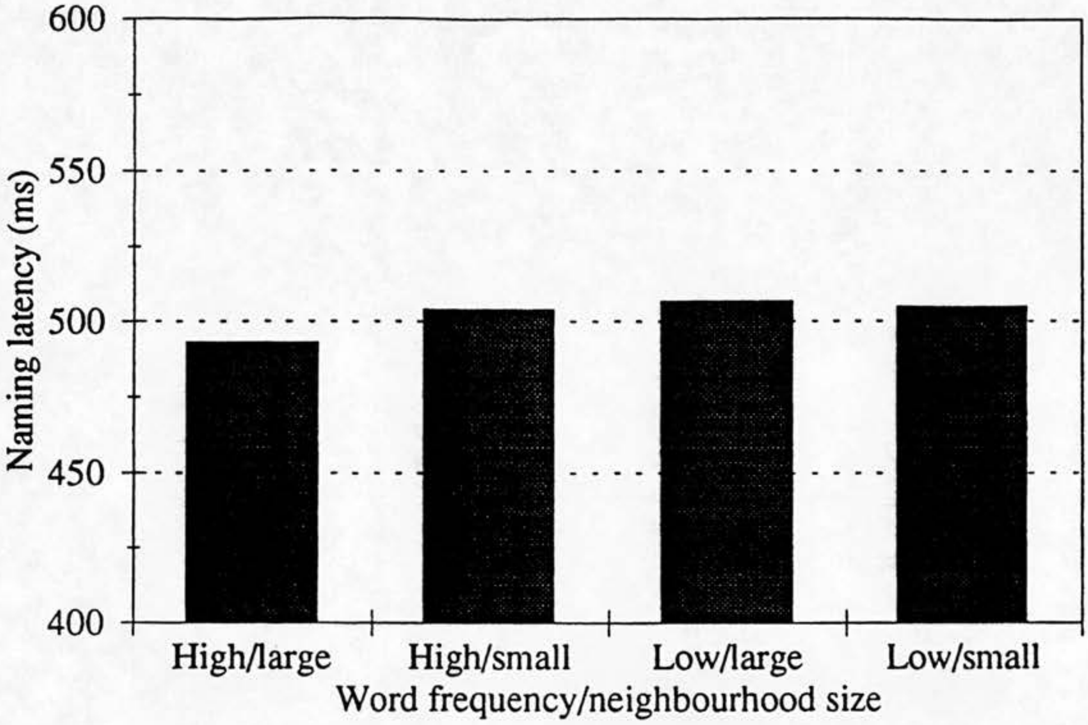

Figure 5–4: Mean naming times (ms) for the word frequency × neighbourhood size conditions of Experiment 4b

$minF'(1,24) = 29.12, p < 0.001$), delayed naming responses were on average 155.5ms faster than immediate naming responses. The effects of word frequency ($F_1(1,22) = 1.18, p = 0.28$; $F_2 < 1$) and neighbourhood size ($F_1(1,22) = 2.95, p = 0.10$; $F_2 < 1$) were not significant.

The only significant interaction effect was between word frequency and naming point, which was significant by $minF'$, ($F_1(1,22) = 12.95$, $p < 0.001$; $F_2(1,36) = 6.56$, $p < 0.014$; $minF'(1,58) = 4.354$, $p < 0.05$). Scheffé tests showed that in immediate naming, frequency facilitates naming times ($F_1(1,22) = 9.6, p < 0.01$; $F_2(1,36) = 6.82, p < 0.05$), whereas in delayed naming frequency has no effect ($F < 1$); $F_2(1,36) = 1.88, p = 0.08$) on naming time. The interaction of word frequency and neighbourhood size (see figure 5–4) was not significant ($F_1(1,22) = 2.00, p < 0.14$; $F_2 < 1$). The error analysis showed no significant results. Similarly there was no interaction between naming point and neighbourhood size ($F_1 < 1; F_2 < 1$), and no three way interaction between naming point, word frequency and neighbourhood size ($F_1(1, 22) = 1.51, p = 0.23; F_2 < 1$).

5.1.5 Summary of Experiment 4

In the LDT task responses were made much more quickly when the non-words were unwordlike than when they were made to look wordlike. Similarly nonwords with large orthographic neighbourhoods took longer to reject than those with small neighbourhoods. Increasing the similarity of word and nonwords makes the discrimination process more difficult and thus LDT responses take longer.

In the naming task responses in the delayed naming condition were given more quickly than responses in the immediate naming condition. The delay period may allow for complete processing of the word to occur before a response is given. In contrast the immediate naming response time will include processing of the stimulus since it is recorded from stimulus onset.

The results showed that high frequency words are responded to more accurately and more quickly than low frequency words in the LDT task. For the immediate naming task word frequency is also facilitatory, but for delayed naming there is no effect of word frequency.

In LDT responses word frequency interacts with neighbourhood size to produce a facilitatory effect of neighbourhood size for low frequency words, but no effect of neighbourhood size for high frequency words. This frequency, neighbourhood interaction was not found for the naming task.

5.2 Connectionist models of the neighbourhood effect

Connectionist models such as the Seidenberg and McClelland (1989) parallel distributed processing model of word recognition have been used to predict, and to explain the effect of orthographic neighbourhood size. The model's architecture consists of activation units joined together by weighted connections. Unlike the Interactive Activation model (McClelland and Rumelhart, 1981) the units or nodes do not correspond to abstract linguistic concepts such as letters, phonemes or words. Instead these concepts are represented by patterns of activity across the distributed units. Experience with words results in modification of the weights associated with connections between the units. Subsequent presentation of a particular word reinstates the pattern associated with that word's previous occurrence. The more often a particular word or word part is presented to the network, the greater its influence over connection weights. In this way the model implicitly encodes the regularities of the language. Since neighbourhood size represents the commonness of an orthographic pattern, the model should be better able to process words with large neighbourhoods than those with less common orthographic patterns.

The model has been implemented in a computer simulation that has been trained on 2897 monosyllabic words. The output of the model takes the form of orthographic and phonological error scores, which are combined into a performance score which is taken to reflect naming performance. Seidenberg and McClelland (1989) have conducted a comparison between the model's performance and the naming latencies obtained by Andrews (1989). The model is reported to demonstrate a "good fit" between the predicted and observed naming data.

	High frequency		Low frequency	
Nhood size	Large	Small	Large	Small
Orthographic score	2.13	2.80	3.01	4.07
Phonological score	2.12	2.23	2.58	4.87
Performance score	4.25	4.35	5.60	9.95

Table 5–5: Simulation 1: Output scores from the Seidenberg and McClelland (1989) model for the stimuli of used in Experiment 4

Since the naming experiment results of Experiment 4b of the present thesis stand in contrast to those reported by Andrews (1989, 1992), the words used in our Experiment 4b were tested on the instantiated simulation of the Seidenberg and McClelland (1989) model.

5.2.1 Simulation 1

As already stated the model produces orthographic and phonological error scores, and a performance score which combines the two. As can be seen from table 5–5 all three output measures produce a similar pattern of results.

A 2 (word frequency: high/low) $\times$ 2 (neighbourhood size: large/small) analysis of variance showed significant effects of word frequency ($F_2(1,29) = 44.04, p < 0.001$), neighbourhood size ($F_2(1,29) = 7.4, p < 0.01$) and a significant interaction between the two variables ($F(1,29) = 5.67, p < 0.02$), for all three types of output score. Figure 5–5 shows the performance score means.

The model produces results which fit well with the lexical decision findings

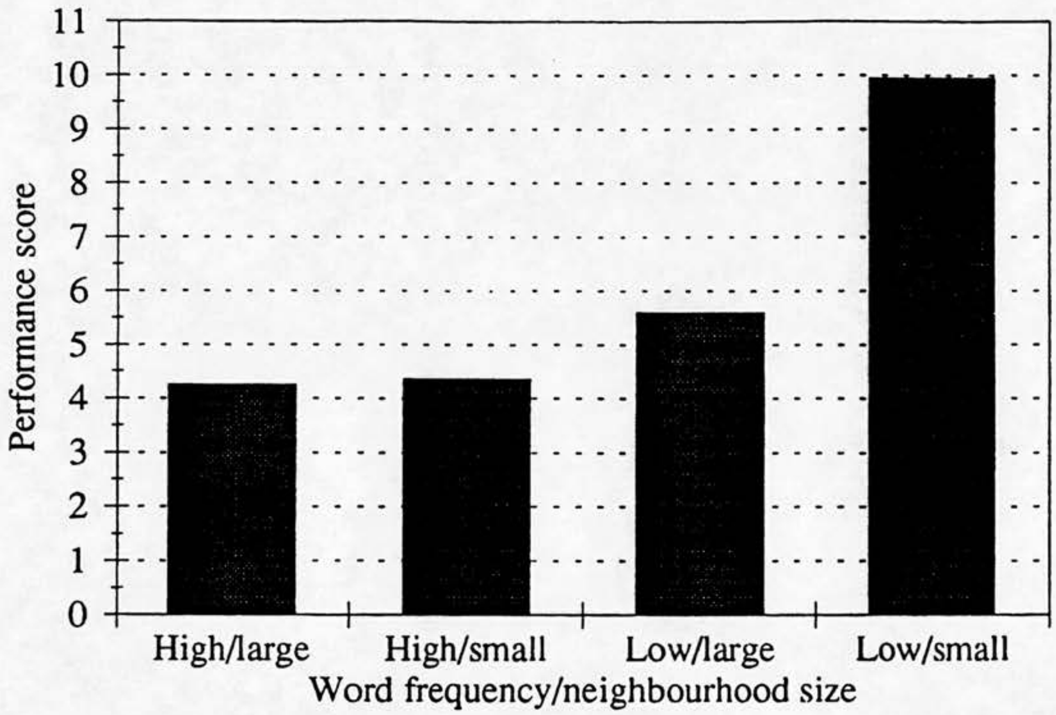

Figure 5–5: Performance scores from the Seidenberg and McClelland (1989) model for the stimuli of Experiment 4

of Experiment 4a, and also with the findings of Andrews (1989, 1992), but not with those reported for naming in Experiment 4b.

5.3 Orthographic neighbourhood size effects in reading

Experiment 4 of this Chapter finds support for Andrews (1989, 1992) finding, and the finding of the Seidenberg and McClelland (1989) model, that large neighbourhoods can be facilitatory for low frequency words. It does not however find support for Andrews' finding that neighbourhood size affects both LDT and naming.

Although Andrews reports neighbourhood effects for LDT and immediate naming, the effect differs between the two tasks. In LDT Andrews reports neighbourhood size facilitation for low frequency words, but in naming neighbourhood size facilitated both high and low frequency words. The neighbourhood size effect for naming was weaker than that for LDT, being significant only in the by-subjects analysis (Andrews 1992 Experiment 2). In Experiment 4 we have found that neighbourhood size was facilitatory only in LDT and not in naming.

Is orthographic neighbourhood size important only in LDT, or is it relevant to a general lexical access process, though highlighted in LDT and overshadowed in naming? Neighbourhood effects could be enhanced in LDT because neighbourhood size is useful to the word/nonword discrimination process. Words which have large neighbourhoods have a common orthographic pattern and therefore look more wordlike than those with a rarer orthographic pattern. Neighbourhood effects could be hidden in naming because they are tangential to a later pronunciation process.

In the following experiment we investigate the effects of neighbourhood size on a reading task which involves lexical access but is neither a decision process nor an overt pronunciation task. Eyetracking was used to monitor fixation times across sentences which contained target words that were manipulated for word frequency, neighbourhood size and concreteness. If orthographic neighbourhood size is important to the general lexical access process, and not just the discrimination process of LDT, then we would expect it to have some influence in reading.

The literature suggests that high frequency words and highly concrete words are fixated for shorter durations than rarer or more abstract words. We expect to obtain similar word frequency and concreteness findings. The literature also suggests that any parafoveal processing which occurs is of a physically specified nature rather than syntactic or semantic (Balota, Pollatsek and Rayner 1985; Rayner 1975). We may expect therefore that words with a common visual pattern or large neighbourhood will benefit more from parafoveal processing and so have shorter fixation times than less common small neighbourhood items.

Design

Eye fixation times were recorded for 64 words presented in neutral contexts. The 64 words fell into 4 cells of a 2×2 manipulation of word frequency (high, low) and neighbourhood size (large, small). The words also conformed to a 2×2 manipulation of word frequency and concreteness (concrete, abstract). There were sixteen different sentence contexts, each was presented 4 times to each subject, once with a word from each of the four frequency $\times$ neighbourhood conditions.

Subjects

The subjects for this experiment were twenty undergraduate and postgraduate volunteers, from the University of Glasgow. All the subjects were native English speakers, skilled readers, and had normal or corrected to normal vision.

There was a hill in the middle	
Prior context	There was a ...
Target	...hill ...
Subsequent context	...in the middle

Table 5–6: Experiment 6: Sentence context example

Materials

Thirty-six of the 40 words comprising the materials from Experiment 4b, and 28 words from Andrews (1989) were presented in sixteen neutral context sentences (e.g. see table 5–6). A list of the sentence contexts is presented in Appendix E.

Procedure

The eye-tracking procedure used was exactly that described in Chapter 3.

5.3.1 Results

For each sentence region (prior context, target, subsequent context) the total fixation time was divided by the number of character spaces within the region to produce a measure of mean character fixation time. The mean character fixations were derived so as to allow comparisons to be made between regions of different lengths. Two separate analyses were undertaken, one to consider word frequency and neighbourhood size, the other to look at word frequency and concreteness.

Word frequency and neighbourhood size. Mean character fixation times of the four conditions, for each of the three sentence regions are shown in Table 5–7. The analysis was a 3(sentence region: prior context, target word, subsequent context) $\times$ 2(word frequency: high/low) $\times$ 2(neighbourhood size:

		Neighbourhood size	
Context	Frequency	Large	Small
Prior	High	31.80	29.48
	Low	28.45	30.63
Target	High	51.12	51.27
	Low	49.88	49.67
Subsequent	High	32.07	33.22
	Low	34.53	31.03

Table 5–7: Experiment 5: Mean character fixation times in ms for the three sentence regions

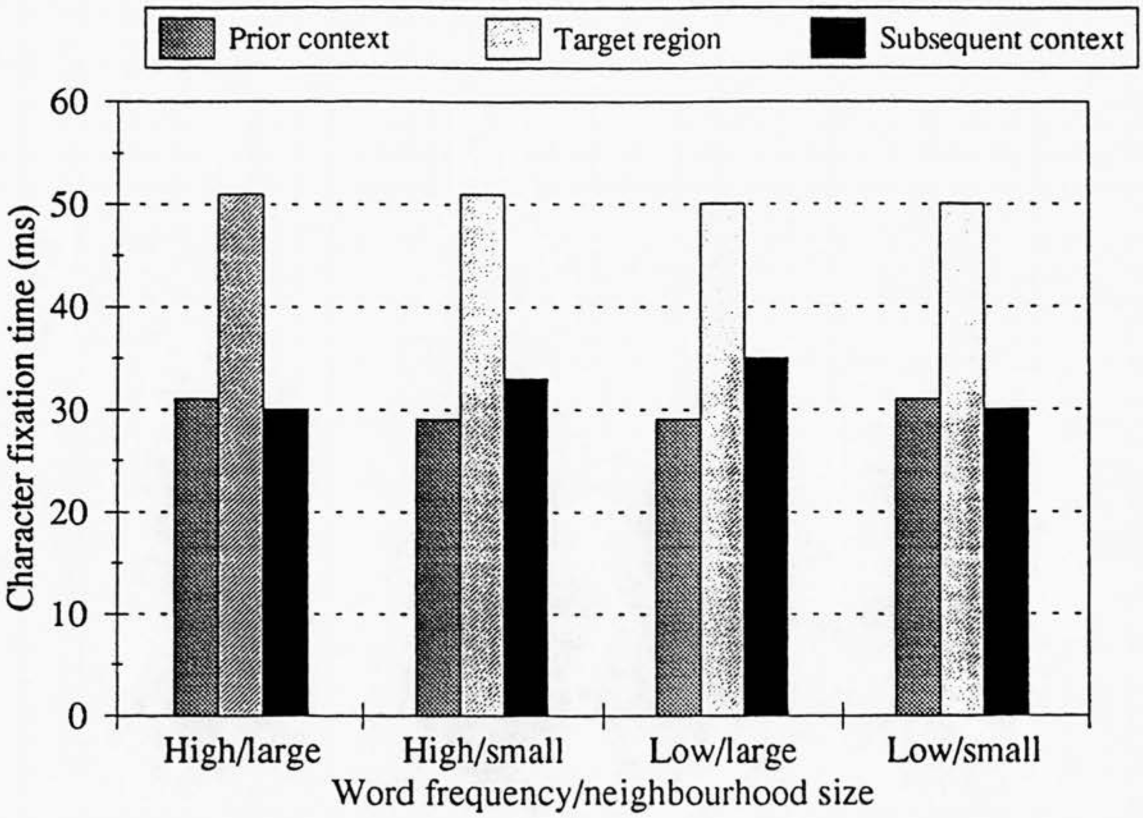

Figure 5–6: Mean character fixation times (ms) for the word frequency × neighbourhood size conditions in each of the three sentence regions.

large/small) analysis of variance. For the by-subjects analysis all factors were within subjects. For the by-items analysis only sentence region was a within items factor, word frequency and neighbourhood size were both between item factors.

The analysis showed an effect of sentence region significant by *min*F', ($F_1(2,38) = 110.04, p < 0.001$; $F_2(2,120) = 219.76, p < 0.001$; $minF'(2,79) = 73.32, p < 0.001$). Scheffé tests showed that target regions were fixated for significantly longer than context regions ($F_1(2,38) = 173.52, p < 0.001$; $F_2(2,120) = 33.06, p < 0.001$).

The main effects of word frequency ($F_1(1,19) = 1.59, p = 0.22$; $F_2 < 1$),

Context	Frequency	Concrete	Abstract
Prior	High	34.75	30.72
	Low	30.46	29.70
Target	High	48.62	51.93
	Low	50.44	49.89
Subsequent	High	33.18	35.25
	Low	32.96	34.73

Table 5–8: Experiment 5: Mean character fixation times for concreteness conditions in the sentence final region

and neighbourhood size ($F_1 < 1, F_2 < 1$), were not significant. None of the two way interactions were significant (all $F < 1$), but the three way interaction between sentence region, word frequency and neighbourhood size (see figure 5–6 was significant in the by subjects analysis ($F_1(2,38) = 4.50$, $p < 0.04$; and marginally significant in the by items analysis $F_2(2,120) = 2.76$, $p < 0.06$).

Separate 2(word frequency: high/low) × 2(neighbourhood size: large/small) ANOVAS were carried out for each of the three sentence regions.

The prior context region showed no effect of word frequency ($F_1(1,19) = 1.15, p = 0.29; F_2 < 1$), and no effect of neighbourhood size ($F_1 < 1; F_2 < 1$). The interaction of word frequency and neighbourhood size was significant in the by subjects analysis, but not the by items analysis ($F_1(1,19) = 5.36, p < 0.03; F_2(1,60) = 2.21, p = 0.14$).

Analysis of the target region showed no effects of word frequency ($F_1(1,19) =$

$1.39, p = 0.25; F_2 < 1$), no effect of neighbourhood size ($F_1 < 1; F_2 < 1$), and no interaction between word frequency and neighbourhood size ($F_1 < 1; F_2 < 1$).

Analysis of the subsequent context region showed no effect of word frequency ($F_1 < 1; F_2 < 1$) or neighbourhood size ($F_1(1,19) = 2.68, p = 0.11; F_2 < 1$). The interaction between word frequency and neighbourhood size was significant by subjects and by items, but not by $minF'$ analysis, ($F_1(1,19) = 9.16, p < 0.006$; $F_2(1,60) = 3.97, p < 0.05; minF'(1,77) = 2.76, p = 0.07$). Scheffé tests showed that regions following high frequency targets were facilitated if that high frequency word had a large neighbourhood ($F_1(1,19) = 16.41, p < 0.01$), and that regions following low frequency words were inhibited if that low frequency word had a large neighbourhood ($F_1(1,19) = 13.00, p < 0.01; F_2(1,60) = 4.00, p < 0.05$).

Word frequency and concreteness. Character fixation time means for all three sentence regions are shown in table 5–8. An overall analysis was conducted using a 3(sentence region: prior context, target word, subsequent context) × 2(word frequency: high/low) × 2(concreteness: concrete/abstract) ANOVA. For the by-subject analysis all factors were analysed within subjects, whereas for the by-items analysis only sentence region was a within items factor, both word frequency and concreteness were between items variables.

The effect of sentence region was significant by $minF'$ analysis ($F_1(2,36) = 79.33, p < 0.001$; $F_2(2,120) = 217.55, p < 0.001$; $minF'(2,64) = 58.13, p < 0.001$). Scheffé tests showed that characters in the target region were fixated for longer than those in context regions ($F_1(2,38) = 132.54, p < 0.001$; $F_2(2,60) = 26.77$, $p < 0.001$). The effects of word frequency ($F_1(1,18) = 1.87, p = 0.18$; $F_2 < 1$) and concreteness ($F_1(1,18) = 1.62, p = 0.21$; $F_2(1,60) = 2.21, p = 0.14$) were not significant. None of the interactions were significant.

Separate 2(word frequency: high/low) × 2(concreteness: concrete/abstract) ANOVAS were carried out for each of the three sentence regions.

The prior context region showed only a marginally significant effect of target word frequency ($F_1(1,18) = 3.88, p = 0.06; F_2 < 1$). Similarly there was only a marginal effect of concreteness for the by subjects analysis ($F_1(1,18) = 3.64, p = 0.07; F_2 < 1$), and no interaction between word frequency and concreteness ($F_1 < 1; F_2 < 1$).

The target word region showed no effect of word frequency ($F_1 < 1; F_2 < 1$), or concreteness ($F_1 < 1; F_2 < 2$), and no interaction between word frequency and concreteness ($F_1 < 1; F_2(1,60) = 1.35, p = 0.24$).

The subsequent context region showed no effect of word frequency ($F_1 < 1; F_2 < 1$), but a significant effect of concreteness in the by subject and by item analysis, but not by $minF'$ analysis ($F_1(1,18) = 4.42, p < 0.04$; $F_2 91,60) = 4.78, p < 0.03$; $minF'(1,53) = 2.31, p = 0.08$): subsequent contexts following concrete words were fixated for less time than those following abstract words. The interaction between word frequency and concreteness was not significant ($F_1 < 1; F_2 < 2$).

5.3.2 Summary for Experiment 5

The eye-tracking study showed that target word regions were fixated for longer than either their prior or subsequent context regions. This is probably because it was the target word alone which provided some meaning or made sense of an otherwise neutral sentence: that is, the target word was the semantic focus of the sentence.

The only sentence region to show any effects of target word frequency, neighbourhood size or concreteness was the subsequent context region. Subsequent contexts following concrete target words had shorter fixation times than those following abstract target words. This may be because if concrete words are processed more quickly than abstract words, then during the processing of concrete words there may be more parafoveal processing of the subsequent context than is possible during the processing of abstract words.

The subsequent context region also showed a word frequency, neighbourhood size interaction. For regions following high frequency words there was a facilitatory effect of target neighbourhood size on subsequent context fixation times. Regions following high frequency, large neighbourhood targets were fixated for less time than regions following high frequency small neighbourhood targets. Conversely for regions that followed low frequency words the effect of target neighbourhood size was inhibitory. Subsequent regions that followed low frequency large neighbourhood targets had longer fixation times than those following low frequency small neighbourhood targets.

5.4 Discussion

At the end of Chapter 4 we concluded that competition processes delay a words recognition and that how words compete depends on the response to be given. The work of Andrews contradicts these conclusions because she reports that neighbours conspire to facilitate recognition, and furthermore she presents an effect of orthographic conspiracy for word naming. In the two experiments of the present Chapter we have demonstrated support for Andrews finding of a neighbourhood conspiracy effect and have shown that orthographic neighbourhood is likely to influence part of a general lexical access process, not just the discrimination process of LDT. If neighbourhood size influences a general lexical access process then it is quite conceivable that neighbourhood size could influence word naming as well as LDT. In Experiment 3 of Chapter 4 and Experiment 4 of the present Chapter, however, we find no effect of neighbourhood size in naming.

In the following discussion we will use the experimental evidence collected in this thesis so far to demonstrate how and why neighbourhood competition effects may be manifest as both competition and conspiracy influences in the recognition process. We will then offer an explanation as to why the neighbourhood effect is unstable in naming while being relatively robust in LDT. Finally we will consider why it is that a connectionist model of word naming reliably produces a neighbourhood effect that we have found only to be robust for LDT.

5.4.1 Time course of activation

In activation models of lexical access the word frequency effect is modeled in one of two ways. Either the activation thresholds of high frequency words are set lower than those of lower frequency words, or the resting activation levels of high frequency words are greater than those of lower frequency words. Whichever of these two descriptions one employs the result is the same, high frequency words will require less evidence to activate them than lower frequency words. Incoming evidence from a visual stimulus will build up over time, so the time taken to accumulate enough evidence to activate a high frequency word will be shorter than the time needed to accumulate enough evidence for a low frequency word to be activated. High frequency words are therefore recognised more quickly than low frequency words.

In both the tasks of LDT and naming we observe the frequency advantage. In delayed naming there is however no advantage for high frequency words, if anything the trend is for low frequency words to be responded to more quickly than high frequency words. The near reversal of frequency effects in delayed naming may reflect the different activation time course of high and low frequency words. In immediate naming a response is required as soon as possible and at this point of processing high frequency words will have become more active than low frequency words. At the later response point of delayed naming, high frequency words may have declined in activation during the delay period, whereas low frequency words may have reached high activation levels by the time of response. Low frequency words may therefore be more active at the point of delayed naming than high frequency words. It may even be that high frequency words have to be reactivated at the point of delayed response.

Ferrand and Grainger (1993) have shown that it is possible to plot independent time courses for the influence of phonological and orthographic information in the early stages of word identification. If a time course can be mapped for orthographic influence then it is possible that the effects of orthographic neighbourhood also have a time course. If this were the case then we would expect high frequency words to show earlier effects of neighbourhood size because they become active more quickly than lower frequency words. Experiments 4 and 5 both showed a facilitatory effect of neighbourhood size. Experiment 4 concerned a LDT task and showed neighbourhood size facilitation for low frequency words. Experiment 5 monitored eye fixations and showed neighbourhood size facilitation for sentence regions following high frequency words. Since LDT involves not just processing the stimulus but also the execution of a response we assume it will reflect a much later stage in processing than that represented by eye fixation times.

To compare the time course of LDT and reading effects it is necessary to consider saccade duration rather than fixation time. The results of Experiment 5 showed that it is the subsequent context of a target word which is affected by the interaction between word frequency and neighbourhood size. Since there was no significant influence of either word frequency or neighbourhood size in the prior context or target word regions we may assume that the processing of the target word is mainly concentrated between the time at which the target is fixated and the time of the subsequent fixation. To measure saccade length we estimated the fixation points of target word regions and their subsequent context region to be approximately in the center of each region. We then calculated the time duration from target fixation, to fixation of the subsequent context. By considering the duration between target and subsequent fixation we have a

		Task	
Word freq	Nhood	Reading	Lexical decision
High	Large	487	522
	Small	495	521
Nhood effect		+8	-1
Low	Large	501	541
	Small	470	568
Nhood effect		-31	+27

Table 5–9: Comparison of LDT latencies from Experiment 4a and reading times from Experiment 5.

measure that represents target word processing time and is thus more comparable to the LDT latencies of Experiment 4 than the originally reported character fixations. These reading time means along with latencies from Experiment 4a are shown in table 5–9.

Consideration of the combined results of Experiments 4 and 5 appears to show that for high frequency words there is a facilitatory effect of neighbourhood size (8ms) early on in processing, whereas for low frequency words this facilitatory neighbourhood size effect (27ms) appears later on during processing. Furthermore the effect of neighbourhood size for low frequency words at the early processing stage reflected in eye fixation, is an inhibitory effect (31ms).

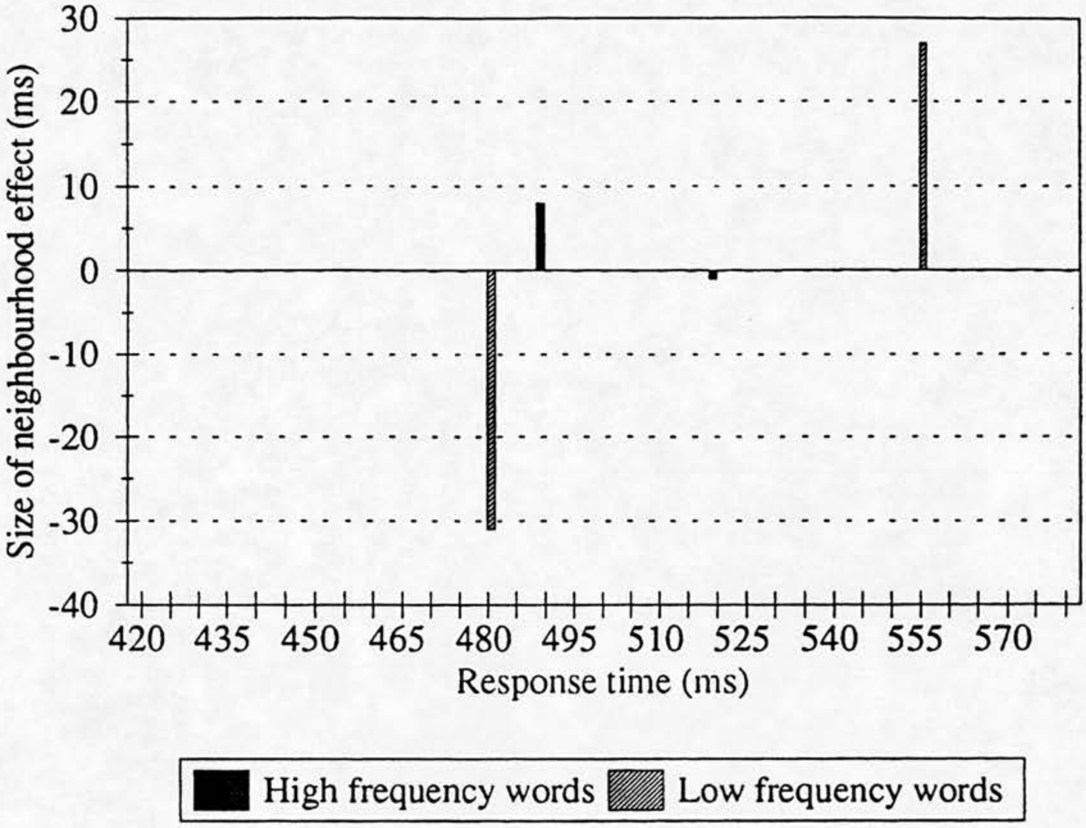

Figure 5–7: Time course of orthographic neighbourhood size effects for high and low frequency words as determined by comparison of Experiments 4 and 5.

At this early stage of processing, low frequency words show inhibitory effects of candidate set size.

As shown in figure 5–7 the influence of neighbourhood size appears to change over the time course of responses. At a very early stage of a word's processing neighbours are competitors (as shown by eye fixations for low frequency words), later during processing they become co-conspirators (as shown by eye fixation times for high frequency words, and LDT latencies for low frequency words) and then later still they are devoid of influence (as shown by high frequency words in LDT). This time course description shows how neighbourhood effects can be both inhibitory or facilitatory at different points in time, but we have not yet explained why an effect should change from being inhibitory to being facilitatory.

The Interactive Activation Model predicts that whole word candidates will compete against one another for recognition. When a word is visually presented the visual feature units become activated and pass activation on to letter units which further activation to word level units. The more common the orthographic pattern of a word, the more words will share its letters and features and thus the more word level representations will be activated. Every activated word unit inhibits all other word units. The stimulus word will therefore inhibit, and receive inhibition from all competing word units. Words with large neighbourhoods should therefore receive more inhibition than small neighbourhood words because they will activate more letter sharing competitors at the word level of representation. For just one word to be recognised, just one word level representation must be active, one word must win out above its competitors. The more active neighbour words a target has, the more competition there will be, and the more slowly the target will be recognised.

One possible mechanism by which the process of competition may reverse to one of conspiracy within an interactive activation model, is that of reverberation. McClelland and Elman (1986) have used the process of reverberation within their TRACE model of speech perception to explain temporal changes in categorical perception. As competitive processes are operating within the system there is feedback from higher to lower levels. This feedback influence becomes stronger over time. When a stimulus is perceived, the influence on lower levels of representation is quite brief since information passes on to higher levels very quickly. As time passes activation feeds back to these lower levels which then have an increasingly important influence over higher levels. This reverberation of activation serves to reinforce or strengthen the evidence for the existence of a words constituent letter and featural parts. A word sharing constituents with many other words therefore receives more reverberating activation than words consisting of more unusual letters or features. A word with many orthographic neighbours would therefore benefit more from reverberation than one with fewer conspiring neighbours.

An interactive activation framework can therefore model a time course of competitor effects which is initially competitive, but which, via a process of reverberating feedback over time, becomes a conspiracy effect. This is exactly what we found in our comparison of Experiments 4 and 5.

5.4.2 Task differences

The experiments of this thesis have shown that neighbourhood size may influence a general lexical access process. The preceding discussion has suggested that the influence of neighbourhood size follows a time course dependent on target word frequency. The neighbourhood influence is primarily one of competition as neighbours compete for dominance, but is then one of conspiracy as neighbours act as co-conspirators. Lexical access is thought to be the process LDT and naming have in common. It should follow, therefore, that both tasks show neighbourhood effects. While there are reports of neighbourhood effects for the naming task in the literature (Andrews 1989, 1992) only effects for LDT have been demonstrated here. Furthermore the effects reported in the literature for naming are weaker than those for LDT.

Aside from the issue of a neighbourhood effect, the word frequency effect is generally believed to be associated with lexical access. Like neighbourhood effects, frequency effects are also weaker for word naming than for LDT (Frederiksen and Kroll 1976). The weaker effect may be due to the use of both lexical and sub-lexical processes in naming, while lexical processing predominates in LDT. If neighbourhood size is a word level effect like word frequency then maybe it is also less reliable in naming than in LDT because naming responses may sometimes be made by constructing pronunciation by rule and thus without reference to whole word representations.

The Seidenberg and McClelland (1989) model has no explicit lexicon, it is a single route model with no separation of lexical and sublexical processes. It is, however, able to demonstrate reliably both word frequency and neighbourhood size effects (see Simulation 1) which we have claimed are dependent on lexical processing.

One of the major criticisms raised against the Seidenberg and McClelland model is its poor performance on reading nonwords (Besner, Twilley, McCann and Seergobin 1990, Coltheart, Curtis, Atkins and Haller 1993). Besner et al. have shown that 40% of pseudowords presented to the model were read incorrectly, and that the errors produced consisted of phonemes which were completely arbitrary with respect to the orthographic input. Besner et al. suggest that the model's nonword performance is more similar to a phonological dyslexia patient than that of a normal reader. The claim is that the model reads many words on an idiosyncratic whole word basis and fails to abstract an effective spelling-to-sound mapping during its training. Each learned word generates a unique hidden unit response in the network and thus it is suggested that the model's behaviour is best understood by seeing it as a distributed lexicon (Besner, Twilley, McCann and Seergobin 1990; Monsell 1991; Zorzi, Houghton and Butterworth 1995).

Plaut and McClelland (1993) have presented an improved version of the Seidenberg and McClelland (1989) model which while claiming to be a single route model exhibits nonword reading performance similar to humans. The newer model incorporates recurrent connections which give rise to attractors. Attractors are best explained in the context of a high-dimensional state-space, in which the activity of each network unit is plotted along a separate dimension. At any moment during processing the pattern of activity across the entire network corresponds to a specific point within this hyper-space. If a network had just two units then the pattern of activity would correspond to a point in space whose x and y co-ordinates represented the activation levels of the two units. In a network with hundreds of units the point for a particular activation pattern will have hundreds of co-ordinates instead of just two. As units are activated the global pattern of activity within the network changes so that the corresponding point moves around the state-space. The activity eventually stabilises to a point corresponding to a familiar steady state. This point we call an attractor. Around each attractor in the state-space there is a region corresponding to a set of similar patterns known as a basin of attraction. If the network is set into a pattern falling anywhere within this surrounding attractor basin region, activity stabilises back into the familiar attractor pattern.

The patterns of activity corresponding to the attractors in the Plaut and McClelland model are the input patterns of the items the network is trained on,i.e. the phonological forms of the words. The generalisation from words to nonwords is achieved by the development of componential attractors which reflect print to sound sublexical correspondences, divided into onset, nucleus and coda. In contrast the attractors which develop for exception words are found to be less componential or even non-componential. Even though the Plaut and McClelland model has only a single pathway there appears to be at least a functional distinction between lexical processes (non-componential attractors) and sublexical processes (componential attractors).

In a very recent paper Zorzi, Houghton and Butterworth (1995) have presented computational evidence to show that a network model acquires lexical properties by means of intermediate representation, i.e. when a hidden layer is present. They have presented a network which in learning the mapping from orthography to phonology in English develops 2 processing. The network automatically segregates into lexical and sublexical type processes, the regular or sublexical component is performed by direct connections from spelling to sound while exception words are handled in a lexical manner by the development of intermediate representations in a hidden unit pathway. Both the Seidenberg and McClelland (1989) and the Plaut and McClelland (1993) model read via a hidden layer, it is therefore not so surprising that the Seidenberg and McClelland model simulates the lexical effects of word frequency and orthographic neighbourhood size. The findings of Simulation 1 do not therefore compromise our claims that neighbourhood competition and conspiracy are lexical effects which are more robust for LDT than for naming because naming is less reliant on lexical processing than is LDT.

Chapter 6

Stimulus list effects

6.1 Effect of stimulus list context in naming

The experimental evidence presented so far shows that neighbourhood size may influence a general lexical access process. The effect of neighbourhood size is however more robust in LDT than in a naming task. In the previous chapter it was suggested that neighbourhood size has a weaker influence in naming because it is a lexical or word level effect and naming may be performed either lexically or sub-lexically. Whether naming is performed lexically by look up or sub-lexically by rule will depend in part on the nature of the stimuli presented. If all the items presented in a stimulus list were exception words then subjects are likely to use a lexical naming process since rule based pronunciations will result in many regularisation errors. Alternatively if the stimulus list consisted of only regular words, rule based pronunciation would be a useful process. We may therefore predict that word frequency effects will be stronger or more reliable in exception rather than regular word lists. The effects observed in a naming task will reflect the type of processing used. The type of processing used for a naming response should, however, depend on the nature of the stimulus list.

To minimise the possible contribution of stimulus list structure it would be desirable to present every possible word to the same subject under the same experimental conditions. In such a complete experimental situation, any influence of stimulus list structure should reflect the nature of the lexicon and so be more naturalistic than those undertaken with any sub-set of lexical items. To present all 50,000 items of an adult English speaker's vocabulary would be impractical, but most experiments involve lists of less than 100 items, representing approximately 0.2 percent of the total lexicon.

6.1.1 The Seidenberg and Waters (1989) Mega Study

Probably the largest single psycholinguistic experiment ever undertaken is the Mega Study of Seidenberg and Waters (1989)[1]. The Mega Study involved a stimulus list of 2977 items, (6 percent of the items contained in the adult vocabulary), and the naming latencies and error rates from 30 US undergraduate subjects. The words presented were monosyllabic and between two and eight letters long. The data were collected over three one hour sessions. The naming latencies were recorded in milliseconds. The error rates are reflected in accuracy scores which are represented as the number of subjects who produced the correct pronunciation of a given target.

In its fully published form, the Mega Study seems certain to prove a very valuable resource to anyone concerned with the mechanisms of word reading. One of its uses in the present situation is to serve as a baseline of naming response latencies which we may assume to be relatively untainted by the confounding effects of stimulus list construction. Radical differences between the Mega Study data and results from smaller studies would seem most likely to be due to the peculiarities of smaller lexical sub-sets.

---

[1]The Mega Study was conducted by Mark Seidenberg and Gloria Waters with an award from the National Institute of Mental Health, award number 46566. Although the Mega Study data has yet to be fully reported and published, Mark Seidenberg has very generously given a copy of the Mega Study item means and error rates to the present author.

Method of data comparison

For any naming experiment conducted it is possible to extract the stimulus items from the mega data and compare the naming times obtained with those found in Seidenberg and Waters' larger slice of the lexicon. There are two ways in which we can use the Mega Study data to examine stimulus list effects:

1. By correlating Mega Study response latencies with those obtained in other studies.

2. By repeating the analysis of other experiments on response latencies taken from the Mega Study.

6.1.2 Results

The Mega Study naming times do show significant positive correlations with naming times from many other sources (see Table 6–1). The correlations are, however, between identical words on the same task, and we might therefore have expected even greater correlation. The largest correlation accounts for 53% of the variance, leaving the remaining 47% (and in some cases much more) attributable to between-study factors.

Repeating the statistical analysis of each experiment using Mega Study naming times yields significant results for only five of the thirteen originally significant effects (see Table 6–2). No significant effects of onset DDC (as defined in Chapter4), or neighbourhood size, were found. The effect of word frequency also failed to replicate for six of the thirteen experiments. Mean response latencies and ANOVA results for all of the comparisons are given in Appendix F.

The re-analysis of Andrews (1989) Experiment 3, did show a significant word frequency effect ($F_2(1,50) = 7.45$, $p < 0.008$), with high frequency words named faster than low frequency words. Mega Study naming times for the items of

Source	Independent variable	Data pts	Data %	Correlation r	Variance r^2	t	Replication p
Present thesis							
Exp. 1b	Word frame	23	77%	0.43	18%	2.18	0.040
Exp. 2	Onset diphone	58	91%	0.56	31%	5.08	0.000
Exp. 3b	NNN	21	70%	0.58	34%	3.16	0.005
Exp. 4b	N'hood size	35	88%	0.55	31%	3.85	0.000
Andrews 1989							
Exp. 3	N'hood size	54	90%	0.61	37%	5.58	0.000
Andrews 1992							
Exp. 2	N'hood size	45	94%	0.44	19%	3.25	0.002
Exp. 4	Bigram freq'	79	99%	0.45	20%	4.50	0.000
Exp. 5	Bigram freq'	44	100%	0.72	53%	6.92	0.000
Chumbley and Balota 1984							
Exp. 4		43	30%	0.14	2%	0.91	0.360
Waters and Seidenberg 1985							
Exp. 1	Regularity	72	100%	0.64	41%	6.97	0.000
Exp. 2	Regularity	48	100%	0.61	37%	5.30	0.000
Exp. 3	Regularity	47	98%	0.66	44%	5.94	0.000
Exp. 6	Word freq'	102	98%	0.56	31%	6.76	0.000

Table 6–1: The correlation of Mega Study naming times with naming times for the same items taken from other studies.

Effects that replicate:

Independent variable	Source
Word frequency	Andrews (1989) Exp. 3
	Waters & Seidenberg (1985) Exp. 1, Exp. 3
Word frequency × word class	Waters & Seidenberg (1985) Exp. 1, Exp 3

Effects that do not replicate:

Independent variable	Source
Word frequency	Andrews (1992) Exp. 2, Exp. 4, Exp. 5
	Waters & Seidenberg (1985) Exp. 2, Exp. 6
	Present thesis Exp. 4b
Neighbourhood size	Andrews (1989) Exp. 3
Onset DDC	Present thesis Exp. 3

Table 6–2: The replication of smaller study effects with Mega Study data.

Waters and Seidenberg (1985) replicated the findings for their Experiments 1 and 3. The two experiments both show that high frequency words are named faster than low frequency words, (Exp.1: ($F_2(1,65) = 11.49$, $p < 0.001$), Exp.3: ($F_2(1,43) = 9.20$, $p < 0.004$)) and that regular words are named faster than exceptions (Exp.1: ($F_2(1,65) = 7.36$, $p < 0.001$), Exp.3: ($F_2(1,43) = 14.69$, $p < 0.0004$)). Experiments 1 and 3 also show a frequency, regularity interaction (Exp.1: ($F_2(1,65) = 5.40$, $p < 0.006$), Exp.3: ($F_2(1,43) = 11.41$, $p < 0.001$)), whereby exception words show much greater word frequency effects than regular words in naming. This is the effect one would expect if word frequency is a lexical effect, and exception words rely more on lexical processing than do regular words. Experiments 1 and 3 of Waters and Seidenberg (1985) are those which include strange (irregular pronunciation and uncommon spelling patterns, e.g. aisle, choir, tsar, yacht) exception words. Experiment 2 from the same paper contains only normal exception words and gives no significant results in the Mega Study re-analysis.

6.1.3 Summary

Correlating Mega Study naming times with naming times from 13 other studies showed that the greatest correlation accounted for only 53% of the variance. Considering this was for the same words on the same task we might have expected a greater degree of correlation.

Repeating the analysis of the experiments with Mega Study naming times showed significant effects for only 5 of the 13 originally significant effects. In general Mega Study materials do not replicate naming time effects found in smaller studies.

The only effects which replicate from other studies when Mega Study naming times are used concern word frequency and pronunciation regularity. Frequent words are named faster than rare ones and regular words are named faster than exceptions. Words which have irregular pronunciations and rare spelling patterns show the greatest word frequency effects.

What the comparison of Mega Study naming times with data from other sources shows is that different responses can be given to the same words on the same task if those words appear in different sorts of stimulus lists.

6.2 Effect of Stimulus list context in LDT

The Mega Study comparison shows that naming responses are affected by the nature of the stimulus list. It is likely that the proportion of regular, irregular and exception words in the stimulus list determines the amount of lexical and sub-lexical naming that takes place. Effects of lexical variables such as neighbourhood size will be especially relevant to the lexical naming process and are therefore less robust in naming than in LDT.

If a naming task promotes sublexical pronunciation processes, then the stimulus list will be influential on dimensions such as spelling-sound consistency (Waters and Seidenberg 1986), or word body consistency (Jared, McRae and Seidenberg 1990). Tasks such as lexical decision which rely more on lexical processing may be sensitive to various dimensions of lexical structure. We may predict that the more unusual a stimulus list characteristic is, relative to the nature of normal lexical experience, the more potent it will be as a context influence. For example approximately 7% of lexical entries have P as their start letter, if 100% of a stimulus list consists of words beginning with P then the activation and reactivation of an initial P will be much greater than one would normally experience.

The lexical dimensions of word frequency and concreteness have very different lexical distributions. As can be seen from figure 6–1, the distribution of word frequency is very skewed. The lexicon contains very many low frequency words

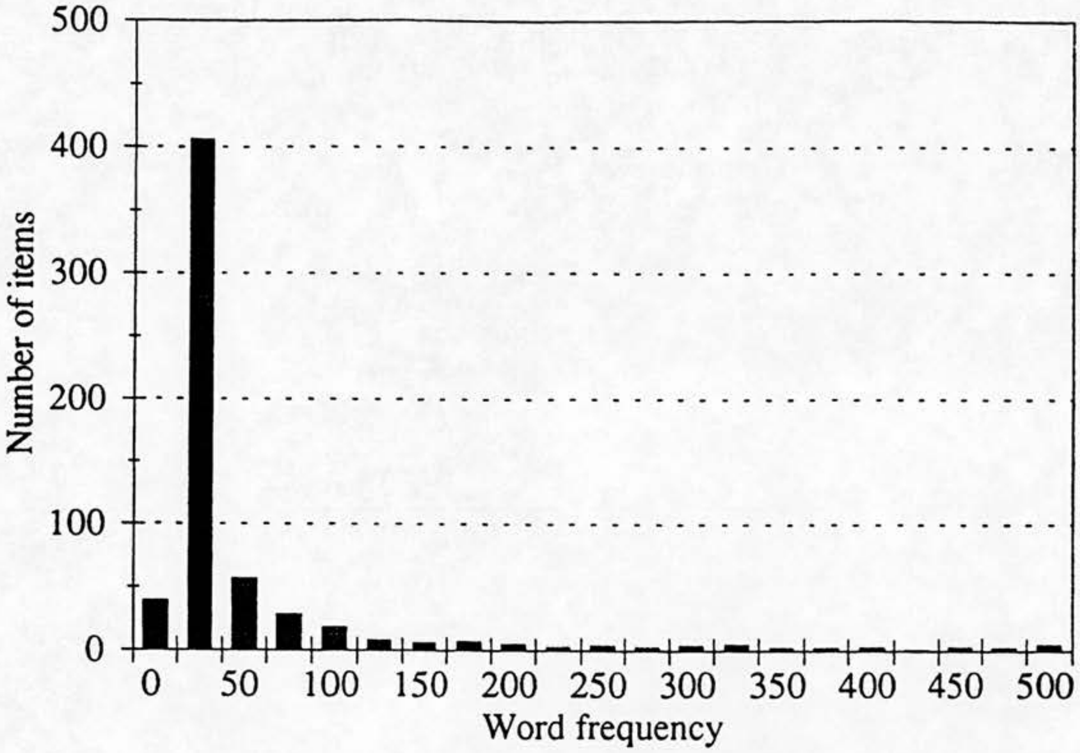

Figure 6–1: Frequency plot to show the distribution of Kučera and Francis (1967) word frequency in a sample of 600 monosyllabic words.

and much fewer very high frequency words. In contrast the distribution of concreteness (as shown in figure 6–2) is far less dramatically skewed.

In the following experiment we investigate the effects of blocking stimulus lists on both word frequency and concreteness dimensions. Since the lexicon consists predominantly of low frequency words, a block of high frequency words should be most unlike normal experience and thus have some contextual impact. Stimuli blocked by concreteness should have less impact on lexical processes because concreteness is more evenly distributed across the lexicon.

In visual word recognition tasks one would hope that any word characteristics correlating with response times would be those directly relevant to lexical access; those features which play a role in natural reading. It is possible however that subjects may be using strategies which are useful to a particular experiment and relevant to a particular stimulus list, but which one would not normally wish

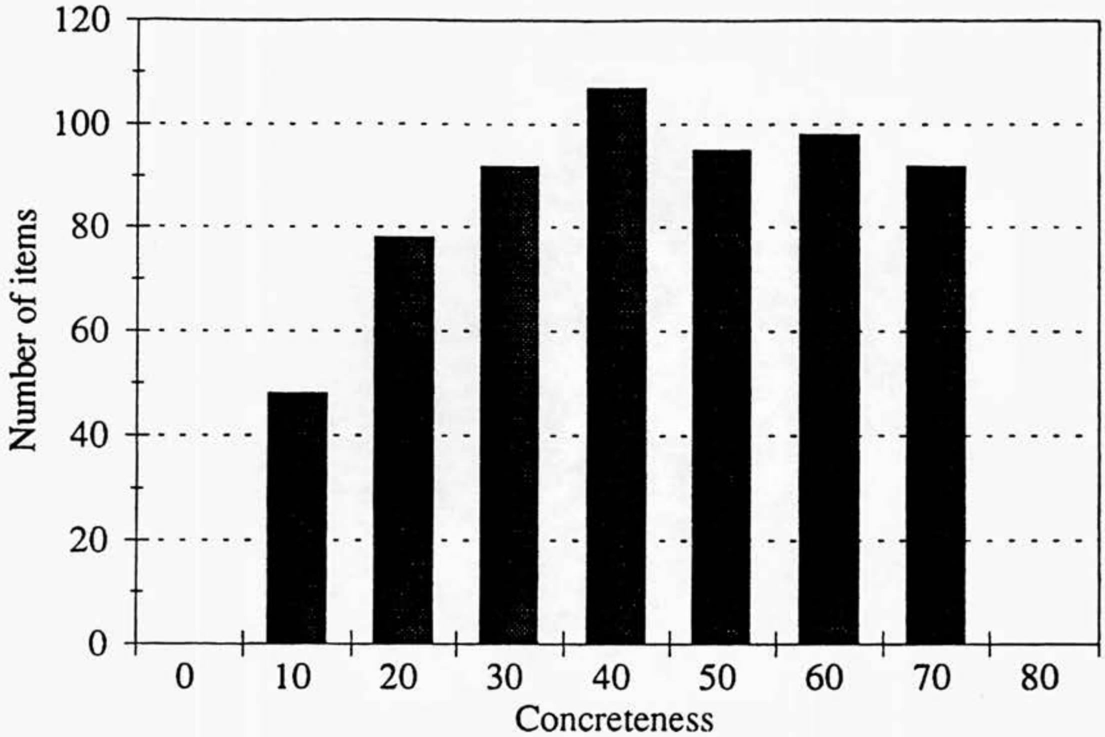

Figure 6–2: Frequency plot to show the distribution of concreteness across a sample of 600 monosyllabic words.

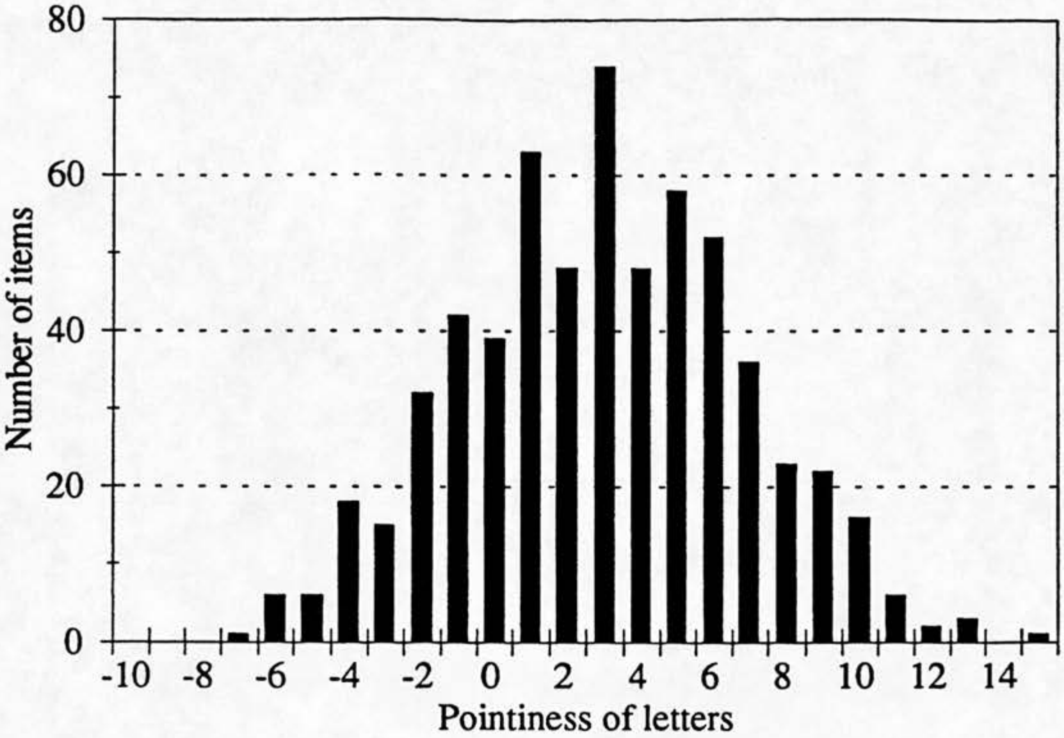

Figure 6–3: Frequency plot to show the distribution of word pointiness across a sample of 600 monosyllabic words.

to claim had any real role in the access process. To investigate the possibility that non-access-relevant stimulus dimensions may be used to determine LDT processing, we examine the effects of a third very different stimulus dimension, which is not normally considered to be relevant to lexical processing. This third dimension is "Word pointiness" a measure of how angular the overall appearance of a word is. Although pointiness is a perceptual dimension one would wish to argue that it should not normally be relevant to lexical processing.

The pattern of word pointiness across the lexicon (as illustrated in figure 6–3) shows a normal distribution, more like concreteness than word frequency. The spread of pointiness like concreteness, means that the probability of encountering a block of round words is roughly the same as that of encountering a block of pointy words. Blocking by pointiness should not therefore be such a potent contextual influence as blocking by word frequency.

If the nature of a stimulus list induces specific top-down influences then a target with the list's characteristic should benefit. Targets antagonistic to the preceding list will not be able to benefit from the top-down contextual input. In the following experiment we examine the effects of providing target words with PRO or ANTI contexts for each of the three above mentioned stimulus dimensions. The effect of PRO or ANTI contexts should show whether or not subjects are being affected by the blocked nature of the stimulus list.

Design

Subjects were assigned to one of four experimental conditions;

- Items blocked by pointiness
- Items blocked by word frequency
- Items blocked by concreteness
- Items blocked randomly

Each target item was presented for LDT after LDT responses had been given to approximately 20 context items which were either PRO or ANTI some dimension of the target word, i.e which were consonant with or different from the target in pointiness, word frequency or concreteness. Whether contexts were PRO or ANTI was manipulated within subjects whereas word dimension (pointiness, frequency or concreteness) was examined using a between subjects manipulation. Each subject gave LDT judgments to all 420 stimulus items which were presented in blocks of 21. Whether the target appeared as the nineteenth, twentieth or the twenty first item varied randomly to decrease the chances of subjects developing expectations concerning the lexical status of the final block item. For each subject half the targets had PRO contexts and half were seen in

ANTI contexts. A control group also saw all 420 items in blocks of twenty one, but the items in these blocks were randomly assigned.

Subjects

The subjects for this experiment were 48 undergraduate and postgraduate volunteers from the University of Edinburgh. All subjects were native English speakers. All subjects had normal or corrected to normal vision.

Materials

Measures of word frequency, orthographic neighbourhood size and concreteness were determined in the same way as for all other experiments in the present thesis. The new measure of pointiness was calculated by firstly giving each lower case Geneva font letter a score (ranging from +20 to −10) for pointiness, where *w* (+20) is very pointy and *o* (−10) is very round. Secondly for each stimulus item the scores of its constituent four letters were added to produce a word score. Words like *case* scored as being very round while those such as *kilt* were very pointy.

Targets: Twenty words were selected which could be divided in three different ways to produce two equally sized sets. The targets were divided either by concreteness to produce a concrete and an abstract set, by frequency to produce a high and a low frequency set, or by pointiness to produce a pointy and a round set. Word frequency, word onset letter and orthographic neighbourhood size were controlled across all six stimulus set types. A list of these target items is given in Appendix G.

Word context items: Two hundred words were selected which could be divided into two groups of 100 either by the concrete/abstract dimension, the high/low frequency dimension or by the pointy/round dimension. Once again word frequency and neighbourhood size were controlled over the four different

		Word dimension			
		Pointiness	Frequency	Concreteness	means
Context	Pro	−45.7	−23.7	+14.2	−18.4
	Anti	−19.0	+5.3	+8.3	−1.7
means		−32.4	−9.1	+11.2	−10.1

Table 6–3: Experiment 7: Mean facilitation (negative values) or inhibition (positive values) effects of context on LDT word responses.

context conditions. To provide a context list for each target, each set of one hundred was divided into ten lists of ten, and across these ten lists word frequency and neighbourhood size were controlled.

Nonword context items: Two hundred nonword foils were constructed so as to be orthographically and phonologically legal. All nonwords were low in pointiness, and had no concreteness or frequency value. They would always be ANTI whatever characterised their block.

Procedure

The procedure followed the general LDT procedure described in Chapter 3.

Between each block of 21 items there was a compulsory 20 second interval followed by the opportunity for further rest if the subject so desired.

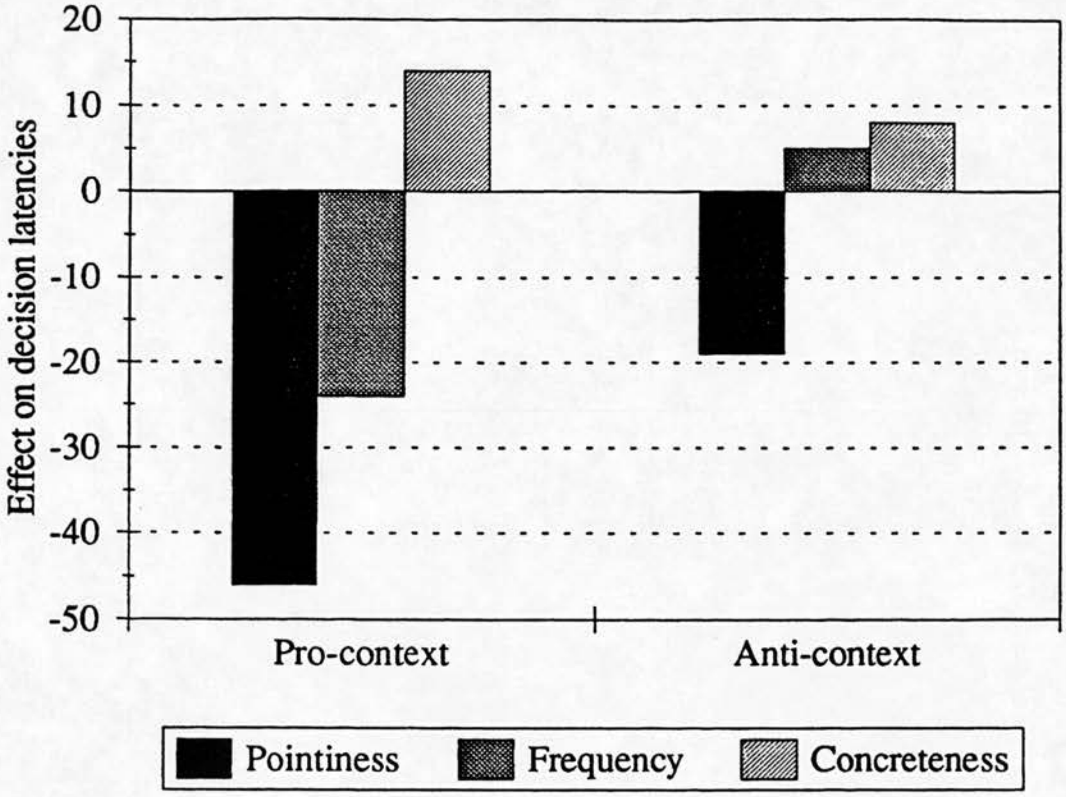

Figure 6–4: Effect of stimuli blocking on lexical decision responses

6.2.1 Results

Response latencies were trimmed using a two standard deviation from the mean cut off as described for Experiment 1. The mean latencies for items in the random context group provided a response time baseline. Subject response latencies in the other three conditions had the baseline subtracted from them to produce response scores. These scores indicated the size of facilitatory or inhibitory effect of the condition compared to the baseline. A 3 (dimension: pointiness/frequency/concreteness) $\times 2$ (context: PRO/ANTI) analysis of variance was performed on target word response scores using both subject and item means as units of analysis. For the by-subjects analysis, dimension was a between subjects factor while context was a within subjects factor. Similarly for the by-items analysis, dimension was a between items factor whereas context was a within items factor. Analyses of error data were not undertaken as total error rate for all subjects under all conditions on all targets was less than 2%.

The mean response scores are shown in table 6–3. Minus scores are those which are faster than baseline and so indicate facilitation. Plus signs show responses slower than baseline: inhibitory effects. The analysis revealed an effect of context which was significant by subjects and items, but not $minF'$; PRO contexts were more facilitatory than ANTI contexts, ($F_1(1,33) = 7.70$, $p < 0.009$; $F_2(1,19) = 4.04$, $p < 0.05$; $minF'(1,38) = 2.65$, $p < 0.07$). The effect of word dimension attained significance only in the by-items analysis, ($F_1(1,33) = 1.90$, $p < 0.16$; $F_2(2,38) = 9.94$, $p < 0.0003$). There was a by-subjects interaction between context and word dimension, ($F_1(2,33) = 3.56$, $p < 0.04$; $F_2(2,38) = 1.85$, $p < 0.17$). Scheffé tests showed that for both the frequency ($F_1(1,33) = 6.54, p < 0.05$) and pointiness dimensions ($F_1(1,33) = 5.51, p < 0.05$) PRO contexts were more facilitatory than ANTI contexts but context made no difference for the concreteness dimension ($F_1 < 1$), responses were inhibited relative to baseline whether the context was PRO or ANTI concreteness (as shown in figure 6–4).

6.2.2 Summary of Experiment 7

The results of Experiment 7 show that context can significantly affect LDT responses. When real words share a context, responses may be facilitated.

The results show that different responses can be given to the same word depending on the dimension used to block stimulus list items.

The dimension/context interaction shows that while agreement with pointiness or frequency of context affected LDT, the dimension of concreteness was not effective. Subjects appear to have been sensitive to frequency and pointiness blocking, but not to concreteness blocking. Concreteness and word frequency are two dimensions we usually consider to be relevant to word recognition, while pointiness is not.

6.3 Blocked presentation of orthographic neighbourhood size

The evidence presented in this Chapter so far shows that both LDT and naming responses are sensitive to the content of the stimulus list used to elicit those responses. The findings of Experiment 7 suggest that the characteristics of a stimulus list determine the nature of top-down input.

In Chapter 5 it was suggested that reverberation of activation from lower levels of representation up to the word level is the locus of neighbourhood conspiracy effects. Neighbourhood effects have been shown to be more robust in LDT than in naming. It should be possible however, to manipulate the stimulus list so as to increase the influence of neighbourhood size on naming responses. Experiment 7 showed that blocking stimuli by a particular dimension to produce a nontypical lexical distribution increased subjects sensitivity to that dimension. The distribution of orthographic neighbourhood size within the lexicon (as shown in figure 5–1 in chapter 5) shows a similar, though less dramatic, pattern to that of word frequency. The lexicon contains more small neighbourhood, than large neighbourhood words. In the following Experiment we therefore investigated the effects of blocking orthographic neighbourhood size, on LDT and immediate naming responses.

6.3.1 Experiment 8a: LDT

Design

The experimental design was a $2 \times 2 \times 2$ manipulation of word frequency (high/low), neighbourhood size (large/small) and neighbourhood size presentation order. Subjects saw either a block of pure large neighbourhood words then a block of pure small neighbourhood words, or the small followed by the large. The subjects were randomly assigned to just one of these presentation order conditions.

Subjects

The subjects were twenty new individuals from the same population as described for the previous experiment.

Materials

The materials for this experiment were the word and nonword stimuli used in Experiment 4a. In addition there were forty extra practice items, twenty words and twenty nonwords. Half of the extra items had large neighbourhoods and half had small neighbourhoods. A block of neighbourhood size blocked practice items was presented before each block of target items.

Procedure

The general procedure was identical to that described for the LDT in Experiment 4a. Here however, the stimuli were not presented randomly but appeared blocked by neighbourhood size. Before each neighbourhood size block there was a practice block of twenty items with similar neighbourhood sizes to those in the following test block.

Nhood	High freq				Low freq				
	Large		Small		Large		Small		
	RT	PE	RT	PE	RT	PE	RT	PE	Mean RT
Presentation order									
Large / Small	504	5.0	493	6.2	559	6.2	578	1.2	533
Small / Large	529	1.2	523	4.9	613	3.7	593	3.7	564
mean	516	3.1	508	4.9	586	4.9	585	2.4	549

Table 6–4: Experiment 8a: Mean lexical decision times in milliseconds (RT) and percent errors (PE) for subjects word responses

6.3.2 Results

Subject and item response means which were more or less than than two standard deviations from the mean were trimmed as described for Experiment 1. A 2 (presentation order: large neighbourhoods then small neighbourhoods/Small neighbourhoods then large neighbourhoods) $\times$ 2 (word frequency: high/low) $\times$ 2 (neighbourhood size: large/small) analysis of variance was used to examine subject and item response times. For the by-subjects analysis word frequency and neighbourhood size were within subjects factors whereas presentation order was analysed between subjects. For the by-items analysis presentation order, word frequency and neighbourhood size were all between item factors.

Analysis of word response times The means in table 6–4 show that when small neighbourhoods are shown first, response times are 31ms slower overall. This main effect of presentation order was significant by items but not by subjects, ($F_1(1,18) = 1.88$, $p < 0.18$; $F_2(1,72) = 15.17$, $p < 0.0002$).

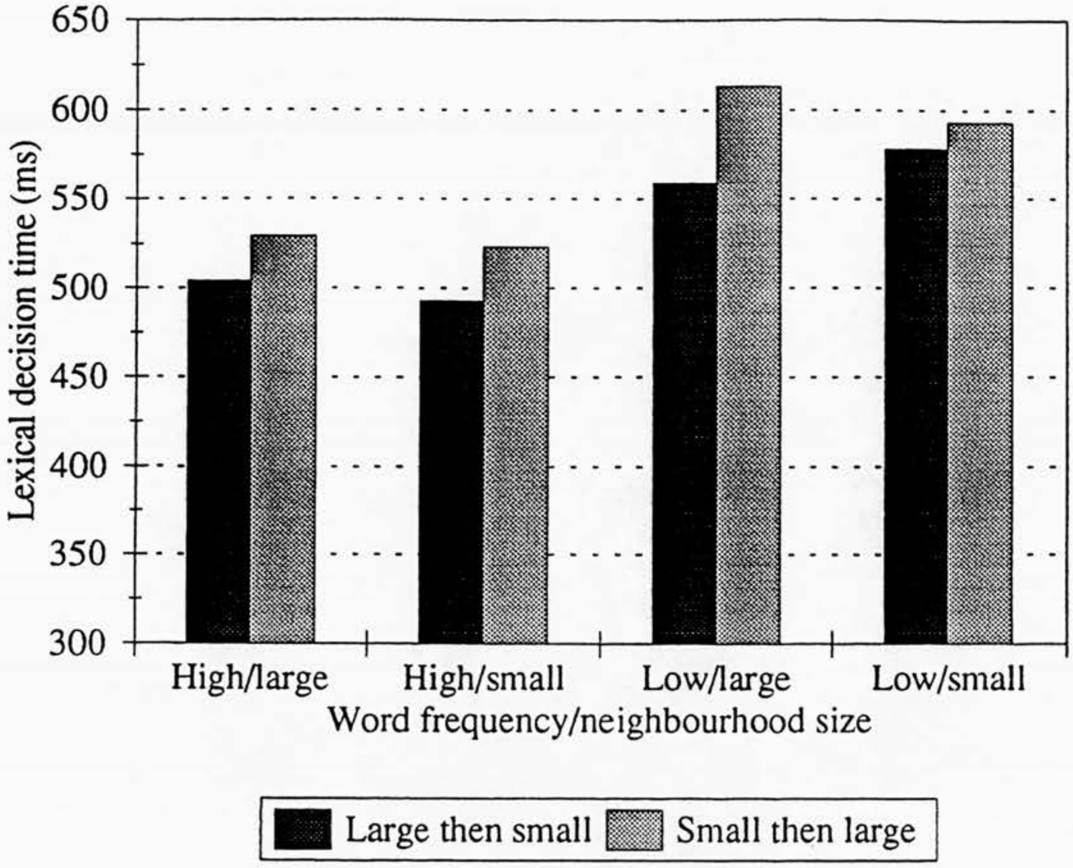

Figure 6–5: Mean lexical decision times (ms) for the word frequency × neighbourhood size conditions of both blocking conditions in Experiment 8a

Analysis of variance revealed a main effect of word frequency that was significant by $minF'$, ($F_1(1,60) = 44.51$, $p < 0.001$; ($F_1(1,18) = 92.38$, $p < 0.00001$; $F_2(1,72) = 85.92$, $p < 0.0001$): high frequency words were responded to 73ms faster than low frequency words.

The neighbourhood size main effect was not significant by subjects or items, ($F_1(1,18) = 0.30$, $p < 0.58$; $F_2(1,72) = 0.33$, $p < 0.56$): The interaction of presentation order and neighbourhood size was not significant ($F_1 < 1$; $F_2(1,72) = 1.07, p = 0.30$).

The three way interaction of word frequency, neighbourhood size, and presentation order (see figure 6–5) was significant by subject analysis only, ($F_1(1,18) =$

	Neighbourhood size				
	Large		Small		
Presentation order	RT	PE	RT	PE	Mean RT
Large / Small	667	4.1	605	2.4	636
Small / Large	675	3.9	643	2.6	659
mean	671	4.0	624	2.5	647

Table 6–5: Experiment 8a: Mean lexical decision times in milliseconds (RT) and percent errors (PE) for subjects nonword responses

4.55, $p < 0.04$; $F_2(1,72) = 2.01$, $p < 0.16$). An examination of this interaction by a Scheffé test revealed that while high frequency words with large neighbourhoods ($F_1(1,18) = 5.41, p < 0.05$), high frequency words with small neighbourhoods ($F_1(1,18) = 8.42, p < 0.01$), and low frequency words with large neighbourhoods ($F_1(1,18) = 25.96, p < 0.001$) were faster when large neighbourhood blocks were seen first, low frequency small neighbourhood words showed no effect of block presentation order ($F_1(1,18) = 1.97, p = 0.09$).

Analysis of nonword response times Analysis of the nonword data showed that large neighbourhood nonwords were responded to significantly more slowly (46.47ms) than small neighbourhood nonwords. This effect was significant by $minF'$, ($F_1(1,18) = 46.00$, $p < 0.00001$; $F_2(1,76) = 23.92$, $p < 0.0001$; $minF'(1,82) = 15.39$ $p < 0.01$).

Subjects who responded to large neighbourhood stimuli before small were 23.15ms slower in responding than subjects who saw small neighbourhood non-

	High freq				Low freq				
Nhood	Large		Small		Large		Small		
Presentation order	RT	PE	RT	PE	RT	PE	RT	PE	Mean RT
Large / Small	500	3.0	512	1.2	502	5.1	518	3.0	508
Small / Large	517	2.9	531	2.6	524	4.3	534	5.0	527
mean	509	3.0	522	1.9	513	4.7	526	4.0	518

Table 6–6: Experiment 8b: Mean naming times in milliseconds (RT) and percent errors (PE) with blocked presentation of neighbourhood

words first. This effect of presentation order was significant by items but not by subjects, ($F_1(1,14) = 0.39$, $p < 0.54$); ($F_2(1,76) = 5.94$, $p < 0.01$).

Analysis of error data Analysis of error data showed that there was no significant difference in the number of errors made under different presentation order conditions, ($t = 0.71$, $df = 18$).

6.3.3 Experiment 8b: Immediate naming

Design and Materials

The design and materials used for this experiment are the same as those used for Experiment 8a. In the present experiment, however, only the word stimuli were used.

Subjects

The subjects were twenty new individuals from the same population as described for the previous experiment.

Procedure

The presentation procedure was the same as for the previous LDT Experiment, but an immediate naming response (as described in Chapter 3) was required.

6.3.4 Results

The effects of word frequency ($F_1(1,18) = 2.51$, $p < 0.13$; $F_2 < 1$) and neighbourhood size ($F_1(1,18) = 2.57$, $p < 0.12$; $F_2(1,72) = 2.854$, $p < 0.09$) were not significant. The effect of presentation order was significant in the by-items analysis, ($F_1(1,18) = 1.11$, $p < 0.30$; $F_2(1,72) = 5.82$, $p < 0.01$). When large neighbourhoods were presented first naming times were 13ms faster than when small neighbourhood stimuli were seen first.

None of the interaction effects were significant.

6.3.5 Summary of Experiment 8

The effects of word frequency, and of nonword neighbourhood size found in Experiment 4 were replicated in Experiment 8. High frequency words were responded to more quickly than low frequency words. Nonwords with large neighbourhoods were rejected more slowly than nonwords with small neighbourhoods.

In both the LDT and naming tasks of Experiment 8 response latencies were faster if the large neighbourhood size block was seen before the block of smaller neighbourhood size targets. For LDT this effect of block presentation order only

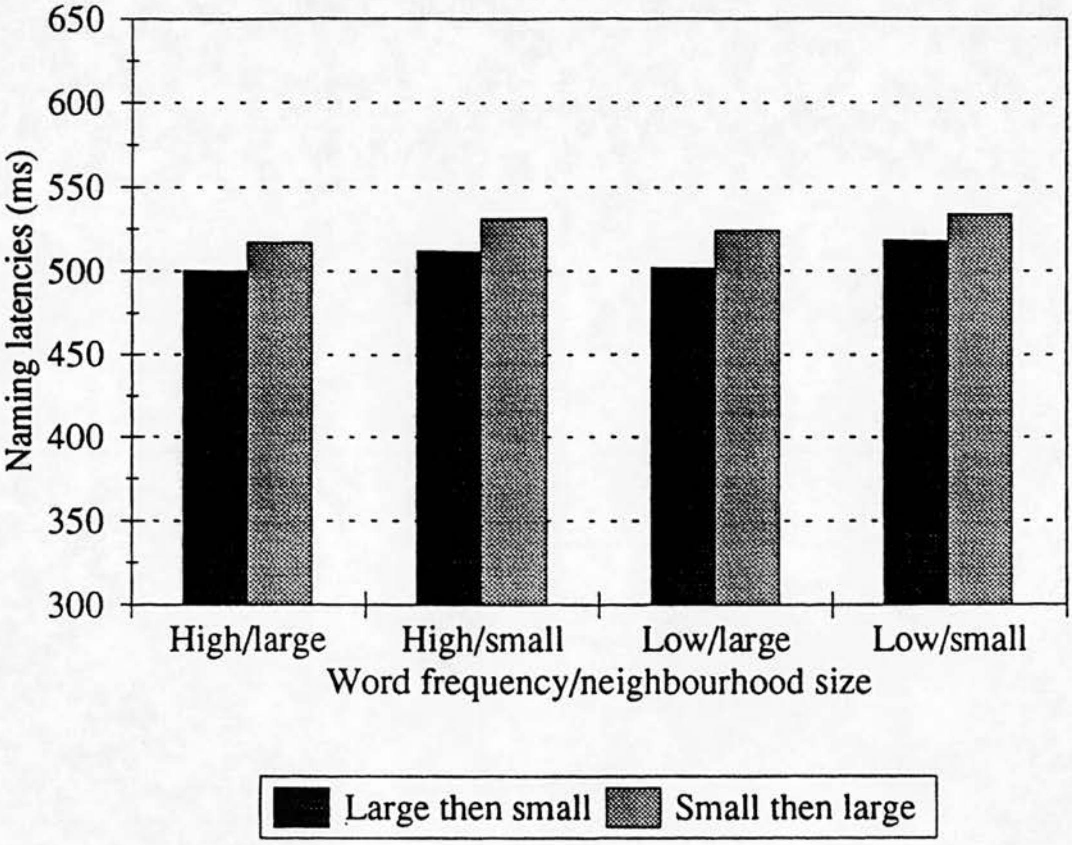

Figure 6–6: Mean naming latencies (ms) for the word frequency × neighbourhood size conditions of both blocking conditions in Experiment 8b

held for high frequency words and low frequency words with large neighbourhoods. Low frequency words with small neighbourhoods were not affected by the block presentation order. Responses to nonwords were slower when large neighbourhood size targets were presented first.

The finding that both LDT and naming responses were influenced by the blocking of orthographic neighbourhood size gives support to the notion that neighbourhood size influences a general lexical access process.

6.4 Discussion

The experimental evidence presented in this Chapter shows that both LDT and naming responses can be influenced by the nature of the stimulus list used to elicit the response. Responses to the same words, on the same tasks were shown to differ across different experiments (see Mega Study comparisons), or even within an experiment across different experimental conditions (see Experiment 7). A LDT or a naming response to a particular word depends not only on the characteristics of that particular word but also on the characteristics of other items in the stimulus list. In the following discussion we will consider how these stimulus list influences may occur, what they influence, and which particular list dimensions will be influential.

6.4.1 Stimulus list influence in word recognition

In naming we have observed that the dimension of the stimulus list which most reliably influences responses is spelling-sound consistency. Regular words are effectively dealt with by grapheme phoneme correspondences but the appearance of exception words in a stimulus list encourages lexical processing. The influence of stimulus list in the naming situation may be to bias the processing in quite a straightforward manner. If words are being pronounced by rule, then an exception word will be assigned the wrong pronunciation and thus processing may revert to a lexical routine.

In the discussion of chapter 5 we considered that rule based processing may be represented as a two layered system of a grapheme to phoneme type mapping. In contrast we suggested that lexical processing has information intermediate to inputs and outputs, and is better represented by multi-layer systems such as TRACE, or the models of Seidenberg and McClelland, or Plaut and Shallice.

Stimulus list effects in a lexical processing system will be more complicated than those influencing a sublexical pronunciation.

In Chapter 5 we suggested that the conspiracy effects of orthographic neighbourhood size may be caused by reverberating activation between word parts and whole word representations. Reverberating activation processes may also account for stimulus list effects. Elman and McClelland (1988) have shown how TRACE can account for the speech phenomenon of phoneme restoration by having feedback reverberating from the word to the phoneme level of representation. An active word unit adds its contribution to the net input of a phoneme unit that it supports. The result is a top-down effect, context is influencing the same detectors that are influenced by bottom-up input.

For the present thesis we are not concerned with words providing a context for phonemes, but rather with stimulus lists providing a context for words. Since Elman and McClelland make the suggestion that their model could be extended to support interactions among higher levels of information such as semantic or syntactic input, our consideration of interactions between stimulus list context and lower levels of representation seems reasonable.

The accommodation of stimulus list effects requires that the lexical processing system has some form of higher representation above that of words. The system should certainly have this higher level or levels since not only can we resolve phonemes in spoken words, but we can also fill whole missing words by reference to the semantic and syntactic context from which they are missed. The only context an isolated word can have is the nature of previously presented isolated words. This is therefore the only source of top-down contextual information which will reverberate back down to the word and lower levels of representation. This source of reverberating activation means that the processing of a word is influenced by previous stimulus list items as well as present bottom up stimulus driven input.

6.4.2 Influential stimulus list information

Blocking of high frequency words in Experiment 7 and of large neighbourhood words in Experiment 8 showed that the contextual effects of stimulus list are most influential when the nature of the stimulus list is in some way inconsistent with normal lexical experience. By this we mean that the greater the distance between the distribution of a variable within a stimulus list, and the distribution of that variable across the lexicon, the greater an impact top-down contextual input will exert.

We have suggested that stimulus list has its locus of influence in reverberation between supra-word and lower levels of representation. For a stimulus list which is a true reflection of lexical structure, the resulting reverberation will have very little impact because the reverberation just supports knowledge the system already knows about. When reverberating activation promotes information that does not comply with normal experience then contextual input becomes a very important source of input.

The finding in Experiment 7 that blocking frequency influences responses while blocking concreteness does not, supports the above suggestions because frequency blocks most dramatically defy lexical experience. The effectiveness of word pointiness as a contextual influence however, requires further explanation.

The nonwords of Experiment 7 were orthographically and phonologically legal. Because nonwords have no lexical entries they could be thought of as being like zero frequency words. They are however not truly equivalent to zero frequency words since subjects do have some knowledge of even zero frequency words. While words can be classed as concrete or abstract, a nonword represents nothing so can have neither a concrete nor an abstract value. The fact that a scale of pointiness can accommodate legal nonwords as easily as words makes it quite different to the dimensions of frequency and concreteness. The words in Experiment 7 could be classed as pointy or as round, but nonwords were more likely to be round. Realisation of the word/nonword decision process as a decision based on word letter shapes may therefore have achieved some degree of success. Pointiness may therefore have been a contextual influence on responses because the dimension was useful to the discrimination process.

6.4.3 Stimulus list and neighbourhood effects

In Experiment 8, unlike Experiments 3 and 4, naming responses showed effects of a neighbourhood size manipulation. The factor that differed between Experiment 8 and the earlier Experiments was the order of stimulus presentation. For Experiments 3 and 4 stimuli were randomly presented, whereas in Experiment 8 stimulus presentation was blocked by orthographic neighbourhood size.

The evidence presented in this Chapter shows that the naming response is sensitive to stimulus list structure. It has been suggested that stimulus list influences determine how much lexical and non-lexical naming occurs. Retrieving pronunciations from lexical lookup is believed to be a faster process than building pronunciations by rule. Experiment 8 showed that naming large neighbourhood size stimuli before small neighbourhood blocks facilitated naming times. We may infer from this that the initial experience of very common orthographic patterns encouraged lexical processing.

Presentation of a large neighbourhood block before a small one had a greater impact on responses than presentation of small followed by large neighbourhood sizes. We have suggested that this is because large neighbourhood words are rarer within the lexicon and so represent a more unusual perceptual experience. Stimulus list characteristics which do not reflect normal lexical experience are more potent sources of contextual influence.

The dimension of orthographic neighbourhood size is not directly relevant to pronunciation since large neighbourhood words (e.g. *have/save*) can be irregular just as often as words with fewer neighbours. A general lexical access process therefore seems a more likely site of influence for neighbourhood size. The finding of a neighbourhood size influence for naming as well as LDT and reading gives support to this suggestion.

Chapter 7

Discussion and Conclusions

In this chapter we consider the findings of the previous three experimental chapters. We discuss the implications of the findings for models of word recognition, and for both existing and future experimental research. We also consider the complexities of establishing the influences upon word recognition. Finally we present a summary of what this thesis has contributed to the understanding of visual word recognition.

7.1 Implications for models of isolated word recognition

In the following sections we discuss the implications that the findings of this Thesis have for models of visual word recognition.

7.1.1 Visual encoding and position specificity

Visual word recognition starts with the perception and encoding of a printed word stimulus. In Chapter 2 we suggested that the outer edges of words may have some special status in perception, since they are more visually distinct than intervening letters. Experiment 2 showed that subjects are sensitive to the frequency of these outer letter frames. In contrast to Experiment 2, Experiment 1 investigated competition operating within the intervening letters between a frame. Competition for word middles was shown to operate in a non-position-specific manner. This finding fits with the observation that these middle letters are less visually distinct.

We could infer from these findings that the boundaries of words are more likely to be encoded as a special fixed position unit, and that in contrast the intervening letters are encoding individually. It is unclear from the work of this Thesis whether frames are encoded first and letters slotted in later, or whether processing of all word parts operates in parallel.

7.1.2 Lexical access and competition

Lexical access may be described as the process by which an encoded perceptual stimulus is mapped onto a representation within lexical memory. Frequency is probably the single most important variable in the lexical access process. Frequency can be thought of as determining the levels of rest activation that word representations have. Word frequency therefore reflects the amount of evidence required before a word can be activated, and thus, the nature of a word's activation time course. For a single word to be accessed, or recognised, it must win out over all other word representations receiving activation from the encoded perceptual stimulus. This stage of the access process is where competition is important, the non-target representations receiving activation are competitors. The activation levels of competitors will depend on how much activation they receive from the encoded perceptual stimulus (i.e. how visually similar they are to the target), and on what their resting activation levels are (i.e. word frequency).

The more competition there is, the more slowly a word is recognised. Experiment 2 showed that there was a competition effect of having very common outer letter frames. Experiment 5 (for low frequency targets) showed a competition effect of having many orthographic neighbours. The NNN candidates in Experiment 1 were both visually similar to, and of higher frequency than, the targets. These single word candidates also showed competition effects on target recognition. The discussion in Chapter 4 suggested that since candidate set size and highest frequency competitor are positively correlated, it may not be the number of competitors that is important, but just the strength of the strongest non-target competitor. Larger competitor sets are more likely to include a very strong competitor and the number of candidates therefore relates to the amount of competition.

7.1.3 Reverberation of activation and conspiracy

In Experiments 4a, simulation 1, and Experiment 5 (for high frequency words) candidate set size, as defined by orthographic neighbourhood, was shown to be facilitatory to target LDT or reading latencies. In contrast to the findings described in the previous section, neighbours appear to conspire rather than compete. It has been suggested, in this Thesis, that orthographic neighbourhood has a time course of influence whereby it first gives rise to competition, but later results in conspiracy. The combined competition and conspiracy effects observed in Experiments 4 and 5 have been used to demonstrate this time course. The mechanism of conspiracy is thought to be reverberating activation. An active word representation reactivates its constituent parts which in turn reverberate activation back to the complete word, thereby increasing the evidence for a words' constituent parts. Before reverberation of activation can start to occur, activation will have to have had time to build up within the system. The conspiracy effects which result from the reverberation of activation are therefore always found after the effects of competition.

7.1.4 Processing routines: one or two?

Like the effect of word frequency, the effect of orthographic neighbourhood size is less robust in a naming task than in LDT. The finding of a neighbourhood effect in reading (Experiment 5) and some sensitivity to neighbourhood size in naming (Experiment 8), suggests that the effect is not an artifact of the decision process of LDT, but has some importance to general lexical processing.

Using a dual route type explanation, we might deduce that because naming can be performed lexically or sublexically, but LDT is more often lexical, the frequency and neighbourhood effects are lexical effects and thus more reliable in LDT. Single route connectionist models challenge the need for separate lexical and sublexical processing, and successfully simulate many word recognition effects. The simulation in Chapter 5, in line with reports from Seidenberg and McClelland (1989), showed significant effects of word frequency, and orthographic neighbourhood size. Despite not having a lexicon the Seidenberg and McClelland model simulates effects which we find to be robust only for LDT. As discussed in Chapter 5, recent criticism of multilayer connectionist networks has involved the claim that hidden units represent an implicit distributed lexicon. The networks are therefore very good at simulating lexical effects, but less successful at performance dependent on strict grapheme to phoneme correspondence, such as nonword reading. This nonword reading task is much better performed by more simple two layer networks. Multilayer networks which are successful in simulating nonword effects (e.g. Plaut and McClelland 1993), employ attractor basins. We have argued that the types of attractor basins which develop for regular and exception words are very different. It would appear therefore that two different sorts of processing routines are still necessary to model accurate reading behaviour.

7.1.5 Top-down contextual influences

The two most common word recognition responses are those of LDT and naming, both of which have been used in this Thesis. The experimental evidence presented in Chapter 6 shows how both tasks are sensitive to stimulus list context. For the naming response, the proportion of regular, irregular and exception words in the stimulus list determines how much of the naming will be lexical and how much will be sub-lexical. Only when the naming is lexical will responses be affected by lexical variables such as word frequency and orthographic neighbourhood size. Only when naming is lexical will we therefore be able to observe lexically based competition effects.

We have suggested in this thesis, that the influence of stimulus list over LDT responses concerns the match of lexical and local experience. Experiments 7 and 8 have shown that when the stimulus list is made to be less representative of the lexical distribution of a certain variable, subjects develop a heightened sensitivity to this stimulus variable. The lexical processing system must therefore have some way of registering mismatch between what it already knows about lexical distributions, and what the present input is telling it. A process of reverberating activation is again appealed to as the mechanism for contextual influence, but this time the relevant activation is reverberating between the word and higher levels of information. We have suggested that when reverberating activation is reinforcing a pattern of information conflicting with normal lexical experience, then the stimulus list becomes a very potent source of influence.

Another related source of information which will influence LDT responses concerns the WORD/NONWORD discrimination. Any dimension of the stimulus list which can promote this discrimination may influence responses. If a certain stimulus dimension equates with wordlikeness then this dimension may become the operationalised definition of a WORD. Stimuli high on this dimension will be responded to more quickly than words which have less of the property.

7.2 Implications for other research

If stimulus list context effects are as important to response processes as this thesis suggests, then serious consideration of these effects may help to explain inconsistencies within the literature. Experiments purporting to investigate the same variable may produce different results if they present the target words in different list contexts. Close examination of stimulus lists may reveal what information is likely to be used during processing and thus what effects may result.

7.2.1 Priming paradigms

In a priming paradigm, a target word is preceded by information that is in some way related to the target. Semantic priming precedes the target with a semantically related word (*lion/tiger*), phonological priming with a phonologically related item (*mayn/main*), and orthographic priming with an orthographically related item (*bolt/belt*). The general hypothesis is that words preceded by a related item should be facilitated. There is much controversy however over which relationships are sensitive to priming. Table 7–1 gives some examples of contradictory findings from the priming literature.

The mechanism most commonly thought to be operating during priming is Spreading Activation. An active word representation is thought to send activation out to other related words. The various priming paradigms attempt to map out this activation spread, and thus show how the lexicon is structured. Priming is thought to affect lexical access because a prime, by pre-activating a

Effect	Example	Found by	Not found by
Semantic priming	lion/TIGER	Deheyer, Briand & Dannenbring (1983) Koriat (1981)	
Phonological priming	mayn/MAIN	Perfetti & Bell (1991)	Martin & Jenson (1988)
Mediated semantic priming	lion/STRIPES	McNamara & Altarriba (1983)	DeGroot (1985) Balota & Lorch (1986)
Mediated phonological priming	famb/SHEEP	Rossen (1983)	McNamara & Healy (1988)

Table 7–1: Experimental evidence for the presence or absence of priming effects.

target, lowers the amount of additional evidence needed for that target to reach activation and attain access.

The contradictory priming literature reports, and the demonstrations of stimulus list influences in Chapter 6, may indicate however that access facilitation is not the whole story. The presentation of *tiger* may facilitate responses to *stripes* and therefore show that the two concepts are related in semantic memory. The repeated presentation of such related pairs within a stimulus list may however increase subjects sensitivity to such relationships being exploited in some sort of response strategy.

The normal practice with priming studies is to have 30% related and 70% unrelated prime/target pairs. To see just how representative these proportions are of normal lexical experience, we considered the distribution of semantic and phonological relatedness within the lexicon. From a randomly ordered sample of 1625 monosyllabic words we calculated how often two items, adjacent in the list, were related to one another. Words were considered to be semantically related if they were category co-ordinates (e.g. *cat/dog*), if one were a property of the other (e.g. *car/drive*), if they had a similar meaning (e.g. *vile/worst*), if one were the superordinate of the other (e.g. *bird/crow*), or if they were strongly associated (e.g. *black/mark*). Phonologically related words were those which rhymed (e.g. *quake/shake*), those which came from the same cohort (e.g. *ramp/rank*), or those which had overall similarity not covered by the previous two definitions (e.g. *width/wealth*). All the related pairs found in the survey, and a breakdown of their relationship, is shown in Appendix H.

The main findings of the survey were that adjacent items were semantically related only 2% of the time, and phonologically related in just 0.8% of cases. The disparity between 1 or 2% and 30% is dramatic. This disparity may be great enough for the structure of the stimulus list to influence target responses beyond the effects of intentionally placed prime items.

One of the most robust priming effects is direct semantic priming (*lion/TIGER*).

Koriat (1981) demonstrated that the effect can even be obtained between pairs of words with very low associative strength (0.065 of the Palermo and Jenkins word association norms). When a stimulus presentation has been preceded by the presentation of a semantically related prime, responses are facilitated. In a converse manner, when a prime and target are semantically unrelated (*shoe/LION*) in a semantic priming experiment, then responses are inhibited (Lorch, Balota and Stamm 1986).

In primed LDT when stimulus lists contain semantically related pairs, the fact that a target is semantically related to the prime guarantees that the correct LDT response is "WORD". This guarantee may make semantic relatedness a very potent response influence. If subjects were equating relatedness to "WORD" responses then unrelated pairs would be bound to suffer. In order to respond "WORD" to an unrelated target the subject has to ignore the evidence that contradicts their semantic expectation and respond positively to the target's lexical status. When unrelated pairs are presented in conditions which do not contain related pairs, they may not exhibit inhibition because they do not contradict any expectations. Further support for this idea of a semantic response criterion comes form Deheyer, Briand and Dannenbring (1983) who report that the semantic priming effect increases in strength as the number of related pairs in the stimulus list increases. With more related pairs the strategy becomes more reliable, or the prime target association better learnt, and subjects become more confident.

In the semantic priming paradigm, *LION* could be used to prime *tiger*, or *TIGER* to prime *stripes*. With mediated priming *LION* is used to prime *stripes*. Mediated priming effects should be weaker than direct semantic priming effects because there is a greater semantic distance between the two concepts. While DeGroot (1983) and Balota and Lorch (1986) both report that mediated priming does not occur in LDT, McNamara and Altarriba (1988) report that it does. McNamara and Altarriba demonstrate that it is the nature of the stimulus list and the response techniques it induces that blocked the effect in earlier studies. If a stimulus list contains a mixture of directly related pairs and mediated pairs, then only an effect of direct relationships will be found. If the list contains no directly related pairs then mediated priming can occur. When directly related pairs are present, subjects appear to use a low semantic criterion (or a very local spread of semantic activation), so that mediated relationships are excluded. When mediated pairs are the only type of semantic relation, subjects adopt a strategy of considering a greater expanse of the semantic network. McNamara and Altarriba found that although using appropriate stimulus lists leads to significant priming effects for both direct and mediated pairs, the effect of direct priming is always stronger. The greater the semantic distance between concepts the wider the spread of activation, the fainter the activation trace and thus the smaller the effect on LDT.

Under normal reading conditions subjects read for meaning. A semantic strategy may therefore be very compelling because it has ecological validity, it is closely related to what people use language for. A semantic strategy is therefore such a strong strategy that it may mask any other potential effects. If phonological and semantic primes exist on the same stimulus list and are presented to the same subjects, then subjects may adopt a semantic response strategy. When phonological primes appear the response strategy will be inappropriate to them and they will be less likely to have a priming effect. Accordingly Martin and Jenson (1988) failed to produce any phonological priming when the stimulus list contained both semantically and phonologically related pairs, although the effect of semantic priming was strong.

Rosson (1983) reports a mediated priming effect whereby *FAMB* primes *sheep*. The implication is that activation can feed between semantic and lexical networks. The pronunciation of *FAMB* activates *lamb* in a lexical network and *lamb* activates *sheep* in a semantic network. Yet McNamara and Healy (1988) found that mediated primes of this type have either a null or inhibitory effect.

McNamara and Healy presented semantic pairs, rhyme pairs and mediated pairs and their controls, thus each subject experienced all three types and their relevant controls. As already discussed a semantic relationship appears to be a very powerful one and if subjects pick up on this as an aid to their response strategy any pairing not exhibiting a semantic relationship will be at something of a disadvantage. The rhyme, or mediated pairs will suffer for not being semantically related.

When experimental stimulus lists are examined to uncover potential dimensions of influence some of the contradictions within the priming literature become explainable. Subjects appear to be sensitive to the most potent relationships between prime and target pairs in a stimulus list. Direct semantic priming is a very robust phenomenon and thus we accept that some semantic relationships are very potent. Direct phonological priming is most effective and reliable when semantic pairs are excluded from the stimulus list. In contrast to semantic and phonological effects mediated priming requires two steps to reach the target and thus is more demanding on the processing system. This could make it a weaker effect and a less desirable response strategy. In situations where subjects can adopt more powerful response strategies it is not surprising that mediated priming does not take effect. Mediated priming would seem to be a much more likely effect in situations where it presents the most useful or only response strategy.

7.3 Defining candidate sets

In this Thesis we have shown competition or conspiracy effects for NNNs, word frames, orthographic neighbours and onset DDCs. We may therefore infer that the word recognition system has some sensitivity to each of these dimensions. We would not wish to claim, however, that the system has separate detector units or processors for every different word dimension. Many word dimensions co-vary with one another. For example high frequency words are also those with

	Word length	Word freq	N'hood size	Bigram freq	Onset DDC	Frame freq
Concreteness	-0.09*	-0.36*	0.10*	-0.14*	-0.10*	-0.14*
Frame freq	0.15*	0.23*	0.02	0.44*	0.04	
Onset DDC	0.03	0.16*	-0.14*	-0.01		
Bigram freq	0.01	0.19*	0.11*			
N'hood size	-0.67*	0.06				
Word freq	-0.17*					

Table 7–2: Correlations between lexical variables in a random sample of 600 monosyllabic words (* significant correlation $p < 0.05$).

frequent letter frames and consistent onset DDCs. Large neighbourhood words are also those with common bigrams.

The correlations in table 7–2 represent the inter-correlations between 7 lexical variables (those considered in this Thesis) within a random 600 word sample of monosyllabic words from the MRC Psycholinguistic database. The correlations show how inherently confounded lexical dimensions are. The activated candidate sets which compete and conspire in the lexical access process may, therefore, not be as specifically determined as neighbours. Instead they may represent more general concept of wordlikeness with which neighbourhood size correlates. Neighbourhood size will, however, not be the only correlate, many other word dimensions will also be correlates. Top-down stimulus list contextual information operates by highlighting any wordlikeness correlate which is unusually represented in the stimulus list.

If many of the examined word dimension variables are intercorrelated, then they will share responsibility for response variance. The unique contribution of a single variable, once other correlates have been held constant may be quite small. By looking for the effect of a single variable in experiments where as many other variables as possible are controlled, what we observe are the residual effects of that single variable. We are left to observe only that portion of the variance in which our critical variable correlated uniquely with the response measure. This fact may explain why the effects of word dimensions are often found to be small or unreliable.

7.4 Summary of Thesis findings

In this final section we will briefly review the main findings of this thesis in relation to the questions which arose from the literature review of Chapter 2, and which we hope we have now answered.

7.4.1 Do candidates compete or conspire?

The existing literature appears to provide conflicting evidence concerning the effects of orthographic neighbourhood size. The effect of having many neighbours, is reported to be one of competition in some instances, and one of conspiracy in others. The work presented in this thesis shows that both of these effects do occur and that both can be accommodated within the same time course explanation. We have explained that competition is the earlier of the two effects because conspiracy operates via reverberating activation, and before activation can reverberate it must have time to build up within a system.

7.4.2 How does word frequency influence competition or conspiracy?

The conspiracy effects of orthographic neighbours found by Andrews (1989, 1992), are present only for low frequency words. High frequency words were not influenced by neighbourhood size. In this thesis we have shown, by eyetracking, that neighbourhood size conspiracy can be observed for high frequency words, but also that neighbourhood competition can be found for low frequency words. These findings are accommodated within out time course explanation by suggesting that high and low frequency words have a temporal activation pattern relative to their frequency. We have demonstrated that in LDT neighbourhood conspiracy operates for low frequency words but that any influence of neighbours has already passed for high frequency words. In eyetracking we obtained earlier response latencies and found that neighbours were conspiring for high frequency words but competing for less frequent words.

7.4.3 Why are neighbourhood effects weaker for naming than for LDT?

Although Andrews (1989,1992) reports neighbourhood conspiracy for both LDT and naming tasks, the effects in naming are smaller than those in LDT. Furthermore in our replication of Andrews' experiments we found no evidence to support neighbourhood conspiracy in naming, while the effect was robust in LDT. The fact that we have found neighbourhood effects in eye tracking, and that such effects do occur in naming at all, supports the claim that neighbourhood size is important to general lexical access and not just to the decision process of LDT. We have suggested that naming can be performed lexically or sublexically and only when it is lexical will neighbourhood size be influential. Neighbourhood effects are thus less robust in naming than in LDT or reading which are more dependent on lexical processing. Single routine connectionist models of word recognition simulate the conspiracy effects of orthographic neighbours. We have suggested that counter to the claims made about these architectures, they can only accommodate the whole range of recognition effects by differentiating lexical and sublexical processing.

7.4.4 How can bigrams be independent of neighbourhoods?

Andrews (1992) found neighbourhood size effects independent of bigram effects. Manipulating orthographic neighbourhood size influenced LDT and naming latencies. Manipulating bigram frequency did not influence either task. Since words from large neighbourhoods are those with a common orthographic pattern, they are also those with frequent bigrams. The finding of a neighbourhood effect, independent of bigram frequency, appears to suggest lexical processing independent of sublexical contribution. In this thesis we have demonstrated the inherent intercorrrelations between word dimensions of various descriptions, and discussed the implications of such confounds. We have claimed that the unique contribution of any one variable, once many other correlates are held constant, will be very small. Manipulating bigram frequency once correlates such as neighbourhood size and word frequency are controlled for, may not leave much variance to affect responses.

7.4.5 What determines the influence of stimulus list?

The word recognition literature contains the suggestion that responses in word recognition tasks may be influenced by the nature of the stimulus list presented. It is clear that subjects may use any sort of information available to aid the WORD/NONWORD discrimination in LDT. In naming the proportion of regular, irregular and exception words in the list will determine whether subjects perform pronunciations lexically or sublexically. Further to these two types of influence, we have shown, in this thesis, that the impact of stimulus list influence is dependent on the disparity between lexical knowledge and local information from the present experimental situation. The greater the disparity between these two types of input, the more potent the influence of stimulus list will be. To explain the process, we have suggested a mechanism of top-down contextual influence operating via reverberating activation. We have discussed the implications of such a mechanism for past and future experimental research.

Bibliography

[1] Adams, M. J. (1979). Models of word recognition. *Cognition, 11* 133–176.

[2] Andrews, S. (1989). Frequency and neighbourhood effects on lexical Access: Activation or Search? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 802–814.

[3] Andrews, S. (1992). Frequency and neighbourhood effects on lexical Access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 234–254.

[4] Balota, D. A. & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 340–357.

[5] Balota, D. A. & Lorch, R. F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 336–345.

[6] Balota, D. A. & Rayner, K. (1983). Parafoveal visual information and semantic contextual constraints. *Journal of Experimental Psychology: Human Perception and Psychophysics, 9*, 726–738.

[7] Bard, E. G. (1990). Competition, lateral inhibition and frequency. In Altman, G. (Ed.)*Cognitive models of speech processing: Psycholinguistic and computational perspectives.* Cambridge, MA: MIT Press.

[8] Bard, E. G. & Shillcock, R. C. (1992). Competition effects during lexical access: Chasing Zipfs tail. Research Paper HCRC/RP-36.

[9] Berry, D. C. & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalisable knowledge. *Quarterly Journal of Experimental Psychology, 36A,* 209–231.

[10] Bery, D. C. & Broadbent, D. E. (1988). Interactive tasks and the implicit–explicit distinction. *British Journal of Psychology, 79,* 251–272.

[11] Besner, D. (1984). Specialised processors sub-serving visual word recognition: Evidence for local control. *Canadian Journal of Psychology, 38,* 94–101.

[12] Besner, D. & Humpheries, G. (1991). *Basic processes in reading: Visual word recognition.* Hillsdale, N. J. Hove: Lawrence Erlbaum.

[13] Besner, D., Twilley, L., McCann, R. S. & Seergobon, K. (1990). On the connection between connectionism and data: Are a few words necessary? *Psychological Review, 97,* 432–446.

[14] Biederman, G. B. (1966). The recognition of tachistoscopically presented five-letter words as a function of digram frequency. *Journal of Verbal Behaviour and Verbal Learning, 5,* 208–209.

[15] Blanchard, H. E., Pollatsek, A. & Rayner, k. (1989). The acquisition of parafoveal word information in reading. *Perception and Psychophysics, 46,* 85–94.

[16] Bouma, H. (1971). Visual recognition of isolated lowercase letters. *Vision Research, 11,* 459–474.

[17] Brown, P. & Besner, D. (1987). The assembly of phonology in oral reading, a new model. *Attention and Performance, XII,* 471–489.

[18] Buswell, G. T. (1935). *How people look at pictures.* Chicago: University of Chicago Press.

[19] Carr, T. H. & Dagenbach, D. (1990). Semantic priming and repetition priming from masked words: Evidence for a center–surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16,* 341–350.

[20] Carroll, J.B., Davis, P. & Richman, B. (1971). *The American Heritage Word Frequency Book.* New York: Houghton-Mifflin.

[21] Cattell, J. M. (1885). The inertia of the eye and brain. *Brain, 8,* 295–312.

[22] Cattell, J. M. (1886) The time taken up by cerebral operations. *Mind, 11,* 220–242, 377–392, 524–538. Reprinted in Poffenberger (Ed.) *James McKeen Cattell* New York: Basic Books, 1973.

[23] Chambers, S. M. (1979). Letter and order information in lexical access. *Journal of Verbal Learning and Verbal Behaviour, 18,* 225–241.

[24] Chumbley, J. I. & Balota, D. A. (1984). A words meaning affects the decision in lexical decision. *Memory and Cognition, 12,* 590–606.

[25] Colombo, L. (1986). Activation and inhibition with orthographically similar words. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 226–234.

[26] Colheart, M. (1977). Critical Notice of Gibson, E. J. and Levin, H. "The Psychology of reading". *Quarterly Journal of Experimental Psychology, 29,* 157–167.

[27] Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.) *Strategies of Information Processing.* London and New York: Academic Press.

[28] Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33A* 497–505.

[29] Coltheart, M., Besner, D., Jonasson, J. T. & Davelaar, E. (1979). Phonological encoding in the lexical decision task. *Quarterly Journal of Experimental Psychology, 31,* 489–507.

[30] Coltheart, M., Curtis, B., Aktins, P. & Haller, M. (1993). Models of reading aloud: dual-route and parallel-distributed-processing approaches. *Psychological Review, 100,* 589–608.

[31] Coltheart, M., Davelaar, E., Jonasson, J. T. & Besner, D. (1977). Access to the internal lexicon. In S.Dornic (Ed.), *Attention and Performance VI* (pp. 535–555). Hillsdale, NJ: Erlbaum.

[32] Dagenbach, T. H. & Carr, T. M. (1990). Inhibitory semantic priming of lexical decisions due to failure to retrieve weakly activated codes. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16,* 328–340.

[33] DeGroot, A. M. B. (1984). Primed lexical decision: Combined effects of the proportion of related prime–target pairs and the stimulus–onset asynchrony of prime and target. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 36,* 253–280.

[34] DeGroot, A. M. B. (1985). Word context effects in word naming and lexical decision. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 37,* 281–297.

[35] Denheyer, K., Briand, K. & Dannenbring, G. L. (1983). Strategic factors in a lexical decision task: Evidence for automatic and attention-driven processes. *Memory and Cognition, 11,* 374–381.

[36] Dorfman, D. & Glanzer, M. (1988). List composition effects in lexical decision and recognition memory. *Journal of Memory and Language, 27,* 633–648.

[37] Elman, J. L. & McClelland, J, L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language, 27,* 143–165.

[38] Everatt, J. & Underwood, G. (1992). Parafoveal guidance and priming effects during reading. A special case of the mind being ahead of the eyes. *Consciousness and Cognition, 1,* 186–197.

[39] Ferrand, L. & Grainger, J. (1993). The time course of orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society, 31,* 119–122.

[40] Forster, K. I. (1976). Accessing the mental lexicon. In E. C. T. Walker and R. J. Wales (Ed.) New Approaches to Language Mechanisms. Amsterdam: North Holland.

[41] Forster, K. I. & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behaviour, 12,* 627–635.

[42] Forster, K. I. & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition, 10,* 680–698.

[43] Forster, K. I. & Davis, C. (1991). The density constraint on form-priming in the naming task: Interference effects from a masked prime. *Journal of Memory and Language, 30,* 1–25.

[44] Forster, K. I., Davis, C., Schoknecht, C. & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *Quarterly Journal of Experimental Psychology, 39A,* 211–251.

[45] Frazier, L. (1983). Processing sentence structure. In K. Rayner (Ed.), *Eye Movements in Reading: Perceptual and Language Processes.* New York: Academic Press.

[46] Frederiksen, J. R. & Kroll, J. F. (1976). Spelling and sound: approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 361–379.

[47] Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical similarity and orthography, concreteness and polysemy. *Journal of Experimental Psychology: General, 113,* 256–281.

[48] Glushko, R. J. (1979). The organisation and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 674–691.

[49] Gordon, B. (1983). Lexical access and lexical decisions: Mechanisms of frequency sensitivity. *Journal of Verbal Learning and Verbal Behaviour, 22,* 24–44.

[50] Gough, P., & Cosky, M. (1977). One second of reading again. In Castellan, N., Pisoini, P., & Potts, G. R. (Eds.) *Cognitive theory, Volume 2.* Hillsdale:Erlbaum.

[51] Grainger, J. (1990). Word frequency and neighbourhood frequency effects in lexical decision and naming. *Journal of Memory and Language, 29,* 228–244.

[52] Grainger, J., O'Regan, J. K., Jacobs, A. M. & Segui, J. (1989). On the role of competing word units in visual word recognition. *Perception and Psychophysics, 45,* 189–195.

[53] Grainger, J. & Segui, J. (1990). Neighbourhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies. *Perception and Psychophysics, 47,* 191–198.

[54] Havens, L. L. & Foote, W. E. (1963). The effect of competition on visual duration threshold and its independence of stimulus frequency. *Journal of Experimental Psychology, 65,* 6–11.

[55] Henderson, L. (1982). *Orthography and Word Recognition in Reading.* London & New York: Academic Press.

[56] Henderson, L. (1985). Towards a psychology of morphemes. In A. W. Ellis (Ed.), *Progress in the psychology of language, Vol. 1.* London: Lawrence Erlbaum Associates Limited.

[57] Henderson, L. & Chard, M. J. (1980). The reader's implicit knowledge of orthographic structure. In U. Frith (Ed.) *Cognitive Process in Spelling.* London & New York: Academic Press.

[58] Huey, E. B. (1908). *The Psychology and Pedagogy of Reading.* Reprinted 1968. Cambridge, Mass.: M.I.T. Press.

[59] Humphreys, G. W., Evett, L. J. & Taylor, D. E. (1982). Automatic phonological priming in visual word recognition. *Memory and Cognition, 10,* 575–590.

[60] Inhoff, A. W. (1982). Parafoveal word perception. A further case against semantic pre-processing. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 137–145.

[61] Inhoff, A. W. (1989). Lexical access during eye fixations in reading: Are words access codes used to integrate lexical information across inter-word fixations? *Journal of Memory and Language, 4,* 444–461.

[62] Jacobs, A. M. & Grainger, J. (1991). Automatic letter priming in an alphabetic decision task. *Perception and Psychophysics, 49,* 43–52.

[63] James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 104,* 130–136.

[64] Jared, D., McRae, K. & Seidenberg, M. S. (1990).The basis of consistency effects in word naming. *Journal of Memory and Language, 29* 687–715.

[65] Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence and the lexicon. *Cognitive Psychology, 13,* 278–305.

[66] Jastrzembski, J. E. & Stanners, R. (1975). Multiple word meanings and lexical search speed. *Journal of Verbal Learning and Verbal Behaviour, 14,* 534–537.

[67] Jordan, T. C. (1986). Testing the BOSS hypothesis: Evidence for position-insensitive orthographic priming in the lexical decision task. *Memory and Cognition, 14,* 523–532.

[68] Kay, J. & Marcel, A. J. (1981). One process, not two, in reading aloud: lexical analogies do the work of nonlexical rules. *Quarterly Journal of Experimental Psychology, 33A,* 397–413.

[69] Kiger, J. I. & Glass, A. L. (1983). The facilitation of lexical decisions by a prime occurring after the target. *Memory and Cognition, 11,* 356–365.

[70] Koriat, A. (1989). Semantic facilitation in lexical decision as a function of prime-target association. *Memory and Cognition, 9,* 587–598.

[71] Kučera, H. & Francis, W. N.(1967). *Computational analysis of present-day American English.* Providence: Brown University Press.

[72] Lorch, R. F., Balota, D. A. & Stamm, E. G. (1986). Locus of inhibition effects in the priming of lexical decisions: Prelexical or post access? *Memory and Cognition, 14,* 95–103.

[73] Luce, P. A., Pisoni, D. B. & Goldinger, S. D. (1990). Similarity neighbourhoods of spoken words. In G. Altman (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives.* Cambridge, MA: MIT Press.

[74] Marchbanks, G. & Levin, H. (1965). Cues by which children recognise words. *Journal of Educational Psychology, 56,* 57–61.

[75] Marslen-Wilson, W. (1990). Activation, competition and frequency in lexical access. In G.Altman (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* Cambridge, MA: MIT Press.

[76] Martin, R. C. & Jensen, C. R. (1988). Phonological priming in the lexical decision task: A failure to replicate. *Memory and Cognition, 16,* 505–521.

[77] Mason, M. (1978). The role of spatial redundancy in grapheme recognition: perception or interference? *Journal of Experimental Psychology: Human Perception and Performance, 4,* 662–673.

[78] Massaro, D. W. (1975). *Understanding Language.* New York & London: Academic Press.

[79] Massaro, D. W., Venexky, R. L. & Taylor, G. A. (1979) Orthographic regularity, positional frequency and visual processing of letter strings. *Journal of Experimental Psychology, 108,* 107–124.

[80] McClelland, J. (1976). Preliminary letter identification in the perception of words and nonwords.*Journal of Experimental Psychology: Human Performance and Perception, 2,* 80–91.

[81] McClelland, J. L. & Elman, J. L. (1986). The TRACE model of spoken perception. *Cognitive Psychology, 18,* 1–86.

[82] McClelland, J. L. & Johnston, J. C. (1977). The role of familiar units in the perception of words and nonwords. *Perception and Psychophysics, 22,* 249–261.

[83] McClelland, J. L. & O'Regan, J. K. (1981). Expectations increase the benefit derived from parafoveal visual information in reading. *Journal of Experimental Psychology: Human Performance and Perception, 7,* 634–644.

[84] McClelland, J. L. & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. Psychological Review, 88, 375–407.

[85] McNamara, T. P. & Altarriba, J. (1988). Depth of spreading activation revisited: Semantic mediated priming occurs in lexical decisions. *Journal of Memory and Language, 27,* 545–559.

[86] McNamara, T. P. & Healy, A. F. (1988). Semantic, phonological, and mediated priming in reading and lexical decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 398–409.

[87] McRae, K., Jared, D. & Seidenberg, M. S. (1990).On the roles of frequency and lexical access in word naming. *Journal of Memory and Language, 29* 43–65.

[88] Meyer, D. E. & Gutschera, F. (1975). Orthographic versus phonemic processing of printed words. Paper presented at a meeting of the Psychonomic Society, Denver, Colorado, November.

[89] Monsell, S., Doyle, M. C. & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General, 118,* 43–71.

[90] Morris, R. K., Rayner, K. & Pollatsek, A. (1990). Eye-movement guidance in reading: The role of parafoveal letter and space information. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 268–281.

[91] Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76,* 165–178.

[92] Morton, J. (1970). A functional model for memory. In D. A. Norman (Ed.) *Models of Human Memory.* New York & London: Academic Press.

[93] Morton, J. (1979). Facilitation in word recognition: experiments causing change in the logogen model. In P. A. Kolers, M. Wrolstad & H. Bouma (Ed.) *Processing of Visible Language I.* New York: Plenum.

[94] Morton, J. (1982). Citation Classic. Interaction of information in word recognition. *Current contents / Social and Behavioural Sciences, 8,* p22.

[95] Navon, D. (1977). Forest before trees: the precedence of global features in visual perception. *Cognitive Psychology, 9,* 353–383.

[96] Norris, D. (1985). So the strong theory loses, but are there any winners? *Behavioural and Brain Sciences, 8,* 718–719.

[97] Norris, D. (1994). Shortlist: A hybrid connectionist model of continious speech recognition. *Cognition, 52,* 189–234.

[98] O'Connor, R. E. & Forster, K. I. (1981). Criterion bias and search sequence bias in word recognition. *Memory and Cognition, 9,* 78–92.

[99] O'Regan, K. (1979). Saccade size control in reading: Evidence for the linguistic control hypothesis. *Perception and Psychophysics, 25,* 501–509.

[100] Parkin, A. J. (1982). Phonological recoding in lexical decision: Effects of spelling-to sound regularity depend on how regularity is defined. *Memory and Cognition, 10,* 43–53.

[101] Parkin, A. J. & Underwood, G. (1983). Orthography vs phonological irregularity in lexical decision. *Memory and Cognition, 11,* 351–355.

[102] Patterson, K. E. & Morton, J. (1985). From orthography to phonology: An attempt at an old interpretation. In K. E. Patterson, J . C. Marshall & M. Coltheart, (Eds.), *Surface Dyslexia* London: Lawrence Erlbaum Associates Limited.

[103] Paivio, A. & O'Neill, B.J. (1970). Visual recognition thresholds and dimensions of word meaning. *Perception and Psychophysics, 8,* 273–275.

[104] Paivio, A., Yuille, J. C. & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 words. *Journal of Experimental Psychology Monograph Supplement,* 76, (3, part 2).

[105] Perfetti, C. A. & Bell, L. (1991). Phonemic activation during the first 40ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language, 27,* 213–234.

[106] Pillsbury, W. B. (1897). A study in apperception. *American Journal of Psychology, 8,* 315–393.

[107] Plaut, D. C., & McClelland, J. L. (1993). Generalisation with componential attractors: Word and nonword reading in an attractor network. In, *Proceedings of the 15th annual conference of the cognitive science society.* Hillsdale, NJ: Erlbaum, 824–829.

[108] Popper, K. (1959). *The logic of scientific discovery.* New York: Basic Books.

[109] Posnansky, C. J. & Rayner, K. (1977). Visual feature and response components in a picture-word interference task with beginning and skilled readers. *Journal off Experimental Child Psychology, 24,* 440–460.

[110] Prinzmetal, W., Hoffman, H. & Vest, K. (1991). Automatic processes in word perception, an analysis from illusory conjunctions. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 902–923.

[111] Rayner, K. (1975). The perceptual span and periferal cues in reading. *Cognitive Psychology, 7,* 65–81.

[112] Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85,* 618–660.

[113] Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8,* 21–30.

[114] Rayner, K. & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory and Cognition, 14,* 191–201.

[115] Rayner, K., Inhoff, A. W., Morrison, R. E., Slowiaczek, M. L. & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 167–179.

[116] Rayner, K. & Morrison, R. E. (1981). Eye-movements and identifying words in parafoveal vision. *Bulletin of the Psychonomic Society, 17,* 135–138.

[117] Rayner, K., Well, A. D. & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception and Psychophysics, 27,* 537–544.

[118] Rayner, K., Well, A. D., Pollatsek, A. & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception and Psychophysics, 31,* 537–550.

[119] Reber, A. S. (1967). Implicit learning or artificial grammars. *Journal of Verbal Learning and Verbal Behaviour, 5,* 855–863.

[120] Reber, A. S. (1976). Implicit learning of synthetic languages; The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 88–94.

[121] Reber, A. S. & Allen, R. (1978). Analogy and abstraction strategies in synthetic grammar learning: A functional interpretation. *Cognition, 6,* 189–221.

[122] Reber, A. S., Kassin, S. M., Lewis, S. & Cantor, G. (1980). On the relationship between implicit and explicit modes of learning a complex rule structure. *Journal of Experimental Psychology: Human Learning and Memory,6*, 492–502.

[123] Rice, G. A. & Robinsson, D. O. (1975). The role of bigram frequency in perception of words and nonwords. *Memory and Cognition, 3,* 513–518.

[124] Richards, L. G. (1976). Concreteness as a variable in word recognition. *American Journal of Psychology, 89,* 707–718.

[125] Rosson, M. B. (1983). From SOFA to LOUCH: Lexical contributions to pseudoword pronounciation. *Memory and Cognition, 11,* 152–160.

[126] Rubenstein, H., Garfield, L. & Millikan, J. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behaviour, 9,* 487–494.

[127] Rubinstein, H., Lewis, S. S. & Rubinstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behaviour, 10,* 645–657.

[128] Rumelhart, D. E. & McClelland, J. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89,* 60–94.

[129] Rumelhart, D. E. & McClelland, J. L. (1986). On learning the past tense of English verbs. In J. L. McClelland & D. E. Rumelhart (Eds.), *Paral-*

lel distributed processing: Explorations in the microstructure of cognition. Cambridge, Mass.: MIT Press.

[130] Rumelhart, D. E. & Siple, P. (1974). Process of recognising tachistoscopically presented words. *Psychological Review, 81,* 99–118.

[131] Segui, J. & Grainger, J. (1991). An overview of neighbourhood effects in word recognition. (Chapter draft)

[132] Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition, 19,* 1–30.

[133] Seidenberg, M. S. & McClelland, J. L. (1989). A distributed developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

[134] Seidenberg, M. S. & Waters, G. S. (1989). Reading words aloud: A mega study. *Bulletin of the Psychonomic Society, 27,* p489.

[135] Smith, F. (1969). Familiarity of configuration vs. discriminability of features in the visual identification of words. *Psychonomic Science, 14,* 261–263.

[136] Snodgrass, J. G. & Mintzer, M. (1993). Neighbourhood effects in visual word recognition: Facilitatory or inhibitory? *Memory and Cognition, 21,* 247–266.

[137] Spoehr, K. T. & Smith, E. E. (1975). The role of orthographic and phonotactic rules in perceiving letter patterns. *Journal of Experimental Psychology: Human perception and Performance, 1,* 21–34.

[138] Stevens, A. L. & Rumelhart, D. E. (1975). Errors in reading analysis using an augmented transition network model of grammar. In D. A. Norman & D. E. Rumelhart (Eds.), *Explorations in cognition.* San Francisco: W. H. Freeman.

[139] Stanovich, K. E., & Bauer, D. W. (1978). Experiments on the spelling to sound regularity effect in word recognition. *Memory and Cognition, 6,* 410–415.

[140] Stone, G. O. & Van–Orden, G. C. (1993). Strategic control of processing in word recognition. *Journal of Experimental Psychology; Human Perception & Performance, 19,* 744–774.

[141] Taft, M. (1979). Lexical access via an orthographic code: The Basic Orthographic Syllabic Structure (BOSS). *Journal of Verbal Learning and Verbal Behaviour, 18,* 21–39.

[142] Taft, M. (1982). An alternative to grapheme–phoneme conversion rules. *Memory and Cognition, 10,* 465–474.

[143] Taft, M. (1991) *Reading and the mental lexicon.* Hove: Erlbaum.

[144] Taft, M. & Forster, K. E. (1967). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behaviour, 15,* 607–620.

[145] Taft, M. & Russell, B. (1992). Pseudohomopne naming and the word frequency effect. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 10,* 51–71.

[146] Thorndike, E. L. & Lorge, I. (1944). *The Teacher's Word Book of 30,000 Words.* New York: Columbia University.

[147] Tinker, M. A. (1958). Recent studies of eye movements in reading. Psychological Bulletin, 55, 215–231.

[148] Treiman, R. & Chaftz, J. (1987). "Are there onset and rime like units in printed words?" In M. Coltheart (Ed.), *Attention and Performance, XII,* London: Erlbaum.

[149] Van Orden, G., C. (1987). A ROWS is a ROSE: Spelling sound and reading. *Memory and Cognition, 15,* 181–198.

[150] Venezky, R. L. (1984) The history of reading research. In P. Pearson(Ed.), *Handbook of Reading Research.* New York: Longman.

[151] Waters, G. & Seidenberg, M. S. (1986). Spelling–sound effects in reading: Time–course and decision criteria. *Memory and Cognition, 13,* 557–572.

[152] Whaley, C. P. (1978). Word–Nonword classification time. *Journal of Verbal Learning and verbal Behaviour, 2,* 143–154.

[153] Winnick, W. A. & Kressel, K. (1965). Tachistoscopic recognition thresholds, paired-associate learning and immediate recall as a function of abstractness, concreteness and word frequency. *Journal of Experimental Psychology, 70,* 163–168.

[154] Zorzi, M., Houghton, G. & Butterworth, B. (1995). Two routes or one in reading aloud? A connectionist dual-process model. *Technical report, UCL.RRG.95.1. March 1995.*

Appendices

Appendix A

Stimuli for Experiment 1

Word stimuli		Nonword stimuli	
Target (NNN)	Control	Target (NNN)	Control
bran	boar	bolb	brot
blot	bury	blad	blar
clam	chap	caln	cyrt
clod	cube	caox	ciba
clot	chop	cilp	cuch
grid	glut	prem	pelf
crud	clef	cirb	cogi
garb	germ	glup	glis
prat	pict	gulm	gops
loin	lute	laof	lerp
slat	soot	salp	sheg
silt	sire	solp	sino
warp	wean	wrom	wege
carp	crab	calw	coim
form	felt	flul	flig

Appendix B

Stimuli for Experiment 2

Word stimuli		Nonword stimuli	
Common frame	Rare frame	Common frame	Rare frame
bunk	bump	buve	bomi
font	flea	fuge	filc
hood	herb	hube	hava
jade	jeer	jift	jarm
lobe	loom	lerd	linv
nook	neap	nule	nawp
teak	tomb	thoe	tarq
feud	fuss	fuer	firb
aqua	arab	amae	apak
eave	exam	emut	erep
nape	numb	nond	nang
oboe	ovum	ogin	ormp
vent	verb	vule	vonk
gore	gear	gret	gemi
glee	gala	gaid	golo

Appendix C

Stimuli for Experiment 3

Consistent onset DDCs

High WF/Large N	High WF/Small N	Low WF/Large N	Low WF/Small N
buck	butt	bust	buzz
miss	milk	mill	mitt
must	much	muck	muff
dick	dish	dill	disc
link	limp	lint	lisp
wave	wait	ware	waif
mail	maid	maze	maim
gaze	gain	gape	gait

Non-consistent onset DDCs

High WF/Large N	High WF/Small N	Low WF/Large N	Low WF/Small N
wall	walk	wart	warp
bare	babe	bail	bait
beat	beef	bead	beta
mark	mars	mast	mask
deal	dean	deer	deem
nose	note	node	nova
name	navy	nail	nape
lack	lamp	lash	lamb

Appendix D

Stimuli for Experiments 4 and 8

High Frequency		Low Frequency	
Large N'size	Small N'size	Large N'size	Small N'size
give	goal	gull	glue
list	lady	loot	lion
name	news	nail	neon
pick	plus	poke	plod
ride	rich	rake	riot
save	suit	seal	scan
must	many	mend	moth
hill	half	hack	hulk
fell	free	fade	fuse
dark	down	dent	dial

Appendix E

Stimuli for Experiment 5

1. Some time later the ...... was all that was spoken about.
2. Soon afterwards they all began to ...... everything they could.
3. It was the ...... that I thought about first.
4. It was the ...... that made me think of it.
5. They tried to ...... everything they possibly could.
6. It was the ...... that had to be dealt with.
7. It was the ...... that he finally asked about.
8. This seemed to be a ...... he thought to himself.
9. He tried to ...... it all
10. After a short time the ...... was almost completely destroyed.
11. He looked at the ...... for quite some time.
12. They tried to ...... it without any success.

13. It looked like a ...... they both agreed.

14. There was a ...... in the middle of it all.

15. After a while the ...... had gone as I had anticipated.

16. There was one ...... and that was all.

Appendix F

Mega study comparisons

Andrews (1989) Experiment 3

Word frequency	High		Low	
Neighbourhood size	Large	Small	Large	Small
Original	580	590	594	643
Mega study	547	544	567	578

	Original	Mega study
Word frequency	$F_2(1, 57) = 8.54, p = 0.01$	$F_2(1, 50) = 7.45, p = 0.008$
Neighbourhood size	$F_2(1, 57) = 6.46, p = 0.05$	$F_2 < 1$
Wf × Nhood	$F_2(1, 57) = 1.62, notsig$	$F_2 < 1$

Andrews (1992) Experiment 2

Word frequency	High		Low	
Neighbourhood size	Large	Small	Large	Small
Original	529	541	569	593
Mega study	539	559	577	562

	Original	Mega study
Word frequency	$F_2(1,44) = 20.9, p = 0.01$	$F_2(1,41) = 2.21, p = 0.14$
Neighbourhood size	$F_2(1,44) = 2.51, notsig$	$F_2 < 1$
Wf × Nhood	*notsig*	$F_2(1,41) = 1.60, p = 0.21$

Andrews (1992) Experiment 4

Word frequency	High		Low	
Bigram frequency	High	Low	High	Low
Original	532	527	546	568
Mega study	555	548	559	562

	Original	Mega study
Word frequency	$F_2(1,76) = 10.15, p = 0.01$	$F_2 < 1$
Bigram frequency	*notquoted*	$F_2 < 1$
Wf × Bf	$F_2(1,76) = 1.86, notsig$	$F_2 < 1$

Andrews (1992) Experiment 5

Word frequency	High		Low	
Bigram frequency	High	Low	High	Low
Original	495	490	527	527
Mega study	544	553	546	551

	Original	Mega study
Word frequency	$F_2(1, 40) = 44.62, p = 0.01$	$F_2 < 1$
Bigram frequency	$F_2 < 1$	$F_2 < 1$
Wf × Bf	$F_2 < 1$	$F_2 < 1$

Waters and Siedenberg (1985) Experiment 1

Word class	Regular		Exception		Strange	
Word frequency	High	Low	High	Low	High	Low
Original	526	518	509	559	519	638
Mega study	565	560	559	592	571	656

	Original	Mega study
Word class	$minF'(1,88) = 10.80, p < 0.01$	$F_2(1,65) = 11.49, p < 0.001$
Word frequency	*notsig*	$F_2(2,65) = 7.36, p < 0.001$
Wc × Wf	$minF'(2,76) = 3.66, p < 0.05$	$F_2(2,65) = 5.40, p < 0.001$

Waters and Siedenberg (1985) Experiment 2

Word class	Regular		Exception	
Word frequency	High	Low	High	Low
Original	552	549	538	580
Mega study	564	560	559	592

	Original	Mega study
Word class	$minF'(1,64) = 5.73, p < 0.05$	$F_2(1,44) = 1.15, p < 0.28$
Word frequency	*notsig*	$F_2(1,44) = 1.23, p < 0.27$
Wc × Wf	*notsig*	$F_2(2,65) = 5.40, p < 0.141$

Waters and Siedenberg (1985) Experiment 3

Word class	Regular		Strange	
Word frequency	High	Low	High	Low
Original	565	546	543	646
Mega study	564	560	571	656

	Original	Mega study
Word class	$minF'(1,63) = 4.49, p < 0.05$	$F_2(1,43) = 9.20, p < 0.004$
Word frequency	*notsig*	$F_2(1,43) = 14.69, p < 0.004$
Wc × Wf	$minF'(1,63) = 5.93, p < 0.05$	$F_2(1,43) = 11.41, p < 0.001$

Waters and Siedenberg (1985) Experiment 6

Word frequency	High	Low
Original	481	497
Mega study	551	555

	Original	Mega study
Word frequency	$F_2(1, 100) = 4.25, p < 0.05$	$F < 1$

Present thesis Experiment 3

Onset	Consistent DDC				Non-consistent DDC			
Word frequency	High		Low		High		Low	
Nhood size	Large	Small	Large	Small	Large	Small	Large	Small
Original	537	538	554	547	535	531	533	551
Mega study	458	455	474	469	469	471	480	482

	Original	Mega study
Word frequency	$F_2(1, 50) = 4.91, p = 0.03$	$F_2(1, 50) = 1.92, p = 0.17$
Onset DDC	$F_2(1, 50) = 3.97, p = 0.05$	$F_2 < 1$
Neighbourhood size	$F_2 < 1$	$F_2 < 1$
All interactions	$F_2 < 1$	$F_2 < 1$

Present thesis Experiment 4b

Word frequency	High		Low	
Neighbourhood size	Large	Small	Large	Small
Original	565	576	586	595
Mega study	537	561	548	568

	Original	Mega study
Word frequency	$F_2 < 1$	$F_2 < 1$
Neighbourhood size	$F_2 < 1$	$F_2(1, 31) = 3.19, p < 0.008$
Wf × Nhood	$F_2 < 1$	$F_2 < 1$

Appendix G

Stimuli for Experiment 7

Concrete		Abstract	
Pointy shaped	Round shaped	Pointy shaped	Round shaped
bank	boss	blow	base
claw	cage	cult	cope
milk	moss	mild	mood
park	page	pull	pass
king	coat	kill	code

High frequency		Low frequency	
Pointy shaped	Round shaped	Pointy shaped	Round shaped
king	coat	cult	cope
kill	mood	claw	code
pull	page	mild	cage
park	pass	milk	moss
bank	base	blow	boss

Appendix H

Relatedness within a random sample of 1625 monosyllabic words

A: Words semantically related to the next word (2%)

Category co-ordinates	Properties	Similar meaning	Association
1. crab/swan	1. ear/rat	1. some/such	1. town/slum
2. elm/bush	2. cold/pond	2. shaft/way	2. beat/race
3. toast/pie	3. white/cloud	3. vile/worst	3. old/weak
4. muff/mask	4. speed/car	4. tune/wail	4. live/breed
5. creek/vale	5. car/drive	5. spear/dart	5. mild/frost
6. rake/spear	6. wave/coast	6. east/place	6. dine/serve
7. dove/fawn	7. show/mare		7. pie/feast
8. yak/grouse			8. tug/leg
9. flesh/spine			9. black/mark
			10. stamp/first

B: Words phonologically related to the next word (0.8%)

Rhymes	Cohorts	Overall
1. quake/shake	1. leg/lens	1. broth/birth
2. cache/dash	2. script/scrape	2. live/love
3. white/write	3. ramp/rank	3. width/wealth
4. bait/gait	4. wan/wad	
5. chose/cruise		
6. tray/slay		